

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 4, 2002, 16:17:35 : Search time 71.77 Seconds
(without alignments)
118.689 Million cell updates/sec

Title: US-09-052-089a-4
Perfect score: 1075
Sequence: 1 KTIINKLFPDLAQEEENVLD.....DLQADQETSLRKSDPP 220

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	ID	Description
1	170	15.8	3210	1	CENP_HUMAN
2	169.5	15.8	1957	1	YD86_SCHPO
3	169	15.7	1976	1	MYHA_BOVIN
4	168.5	15.7	1325	1	G160_MOUSE
5	167.5	15.6	1203	1	XCPC_XENLA
6	166	15.4	2017	1	MTSN_DROME
7	163	15.2	1790	1	USO1_YEAST
8	162.5	15.1	2116	1	MYSD_DICDI
9	162	15.1	895	1	RA50_THEVO
10	162	15.1	1938	1	MYA_ABOIR
11	160.5	14.9	1744	1	TANA_XENLA
12	160.5	14.9	1966	1	MTSB_CABEL
13	160	14.9	1972	1	MYHB_MOUSE
14	159	14.8	963	1	KINH_HUMAN
15	159	14.8	1972	1	MYHB_HUMAN
16	159	14.8	1972	1	MYHB_RABIT
17	157.5	14.7	886	1	RA50_ARCEU
18	156.5	14.6	1509	1	MTSN_ACACA
19	155	14.4	962	1	VDP_HUMAN
20	154.5	14.4	2411	1	EP15_DROME
21	154	14.3	896	1	EP15_HUMAN
22	154	14.3	1976	1	MYHA_HUMAN
23	153.5	14.3	1938	1	MYSD_CABEL
24	153	14.2	1093	1	TMFL_HUMAN
25	153	14.2	1938	1	MYHD_HUMAN
26	152	14.1	1969	1	MYSA_CABEL
27	151.5	14.1	995	1	HIP1_HUMAN
28	151	14.0	880	1	RA50_PYPAB
29	150.5	14.0	882	1	MYSP_CABEL
30	150	14.0	976	1	SCPL_HUMAN
31	150	14.0	1937	1	MYH8_HUMAN
32	150	14.0	1961	1	MYH9_RAT
33	149	13.9	1935	1	MYSS_CYPCA

34	149	13.9	1960	1	MYH9_HUMAN	P35579 homo sapien
35	149	13.9	1978	1	MYHB_CHICK	P10587 gallus gall
36	148.5	13.8	501	1	MYSD_RABIT	P09105 oryctolagus
37	148.5	13.8	727	1	MFPI_ARATH	P01485 arabidopsis
38	148.5	13.8	1290	1	XCPC_XENLA	P50532 xenopus lae
39	148.5	13.8	1427	1	REST_HUMAN	P30622 homo sapien
40	148.5	13.8	1934	1	MYH7_MESAU	P13540 mesocricetu
41	148.5	13.8	1935	1	MYH7_HUMAN	P12883 homo sapien
42	148.5	13.8	1935	1	MYH7_PIG	P79293 sus scrofa
43	148	13.8	1102	1	MYSC_CHICK	P29616 gallus gall
44	147.5	13.7	579	1	G160_HUMAN	P08378 homo sapien
45	147.5	13.7	880	1	MYSP_BRUMA	O01202 brugia mala

ALIGNMENTS

RESULT 1
ID CENP_HUMAN STANDARD: PRT: 3210 AA.
AC P49454: Q13246: Q13171:
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE CENP-F kinetochore protein (Centromere protein F) (Mitosis) (AH antigen).
GN CENPF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Breast carcinoma;
RA MEDLINE=95348175; PubMed=7542657;
RA Liao H., Winkfein R.J., Mack G., Ratner J.B., Yen T.J.;
RT "CENP-F is a protein of the nuclear matrix that assembles onto kinetochores at late G2 and is rapidly degraded after mitosis.";
RL J. Cell Biol. 130:507-518(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95379848; PubMed=7651420;
RA Zhu X., Mancini M.A., Chang K.-H., Liu C.-Y., Chen C.-F., Shan B., Jones D., Yang-Feng T.L., Lee W.-H.;
RT "Characterization of a novel 350-kilodalton nuclear phosphoprotein that is specifically involved in mitotic-phase progression.";
RL Mol. Cell. Biol. 15:5017-5029(1995).
RN [3]
RP SEQUENCE OF 2194-3210 FROM N.A.
RX MEDLINE=95336446; PubMed=7612011;
RA Li Q., Ke Y., Kapp J.A., Fertig N., Medsger T.A. Jr., Joshi H.C.;
RT "A novel cell-cycle-dependent 350-kDa nuclear protein: C-terminal domain sufficient for nuclear localization.";
RL Biochem. Biophys. Res. Commun. 212:220-228(1995).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=95370296; PubMed=7642639;
RA Zhu X., Chang K.-H., He D., Mancini M.A., Brinkley W.R., Lee W.-H.;
RT "The C terminus of mitotin is essential for its nuclear localization, centromere/kinetochore targeting, and dimerization.";
RL J. Biol. Chem. 270:19545-19550(1995).
RN [5]
RP CHARACTERIZATION.
RX MEDLINE=96437347; PubMed=9763420;
RA Chan G.K.T., Schaar B.T., Yen T.J.;
RT "Characterization of the kinetochore binding domain of CENP-E reveals interactions with the kinetochore proteins CENP-F and hBUB1.";
RL J. Cell Biol. 143:49-63(1998).
CC -1- FUNCTION: PROBABLY REQUIRED FOR KINETOCORE FUNCTION, INVOLVED IN CHROMOSOME SEGREGATION DURING MITOSIS. INTERACTS WITH RTINOBLASTOMA PROTEIN (RB), CENP-E AND BUB1.
CC -1- SUBUNIT: HOMO- OR HETERODIMER.
CC -1- SUBCELLULAR LOCATION: NUCLEAR MATRIX (BUT NOT IN THE NUCLEOLUS),

```

CC REORGANIZATION TO THE KINETOCORE/CENTROMERE (CORONAL SURFACE OF
CC THE OUTER PLATE) AND THE SPINDLE DURING MITOSIS.
CC -1- DEVELOPMENTAL STAGE: GRADUALLY ACCUMULATES DURING THE CELL CYCLE.
CC -1- PTM: HYPERPHOSPHORYLATED DURING MITOSIS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U19769; AAA82889.1; -
DR EMBL: U30872; AAA82935.1; -
DR HSSP: P02649; ILE4.
DR MIM: 600236; -.
KW Chromosomal protein; Nuclear protein; Centromere; Coiled coil;
FT DOMAIN 14 197
FT DOMAIN 14 197
FT DOMAIN 273 769
FT DOMAIN 823 1328
FT DOMAIN 1642 1746
FT DOMAIN 1862 2987
FT DOMAIN 2207 2568
FT REPEAT 2207 2386
FT REPEAT 2389 2568
FT DOMAIN 3015 3032
FT DOMAIN 16 16
FT CONFLICT 250 250
FT CONFLICT 272 272
FT CONFLICT 611 611
FT CONFLICT 1494 1589
FT CONFLICT 1611 1611
FT CONFLICT 1811 1811
FT CONFLICT 2242 2243
FT CONFLICT 2335 2335
FT CONFLICT 2492 2492
FT CONFLICT 2545 2561
SQ SEQUENCE 3210 AA; 367589 MW; 11D83324960E4334 CRC64;

Query Match 15.8%; Score 170; DB 1; Length 3210;
Best Local Similarity 24.3%; Pred. No. 0.039;
Matches 69; Conservative 51; Mismatches 96; Indels 68; Gaps 9;

QY 1 KTINKLFEDIAEEN---VLDAEFLKNEIDSVKAQL-----SOK 38
DB 2249 KDEVNEIERELQMESENEQVILDAENSKAEVETLKQIEEMARSLKVFELDLVTLRSEK 2308
QY 39 D-----REKDSQAIIIDTLRDTL-----EERNATVESLQNALNKAEMLC 77
DB 2309 ENLTQKQIOEKQGLSELDKLLSKSLSEKEQAEIQIKESKTAVEMLQNLKELNEAV 2368
QY 78 STL---KKOMKFLBOROD---ETKQAREEFAHRLCKMTQETLEILOSORSE----- 124
DB 2369 AALCGDQEIEMKATQSDPRLPEEBHQLNSTIEKRLAREADEKQOLCVLQQLKSEHHAD 2428
QY 125 -----VEEMIRDMGVGSAVQALVAVCYSLKKEYENLKARKATGELADRLKKDLVSSRS 179
DB 2429 LTKRVEVLEIRELRIARNGHAALEAENSKGEVETLKAKLEGMTQSLRGLELDVVTIRS 2488
QY 180 KLTKLNTLELDQ-----AKLEL--RSAOKDLOSADQETLSLRKS 216
DB 2489 EKEDLTNELQKEQERISELETLINSSFENILQEKQOEKVMKES 2532

RESULT 2
YD86_SCHPO STANDARD; PRT; 1957 AA.
AC Q10411;

```

```

DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical 222.8 kDa protein Clf3.06C in chromosome I.
GN SPAC1f3.06C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z70690; CA94624.1; -
DR Hypothetical protein.
SQ SEQUENCE 1957 AA; 222785 MW; 3F480CA06171D9DA CRC64;

```

```

Query Match 15.8%; Score 169.5; DB 1; Length 1957;
Best Local Similarity 22.4%; Pred. No. 0.025;
Matches 54; Conservative 54; Mismatches 96; Indels 37; Gaps 4;

QY 15 EENVLDAEFLKNEIDSVKAQLSQDKREKDSQAIIIDTLRPLEERNATVESLQNALNKAE 74
DB 1386 EDNOLATNKKKLNQDLHNLQETRLKEDVLEKEKESLISLESLSNQROKESILDAKHELE 1445
QY 75 -MLSTLKQKPFLEORO-----DETQAREEFAHRLCKMTQETLEILOSORSEVE 126
DB 1446 HMLDDTSKSNLSMEKIESINSSLDKSFELASAVEKIGALQKLHSESLMEKINSOLO 1505
QY 127 EMIRDMGVGSAVQALVAVCYSLKKEYEN-----LKEARRATG 164
DB 1506 EAKKEIGVDESTIOELDHEITASKNNEGKLNKDSIIRDSENTEQLNNLAEKSAVK 1565
QY 165 ELADRLKKDLVSSRSKLT-----NTELDQAKLELRSAQKLOSDQETLSLRKSD 217
DB 1566 RLSTERESELQFNSRLADLEYHKRSQVESELGRSKLKLASTTEBLQLAENERLSLTTRML 1625
QY 218 D 218
DB 1626 D 1626

RESULT 3
MYHA_BOVIN STANDARD; PRT; 1976 AA.
AC Q27991;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, nonmuscle type B (cellular myosin heavy chain,
DE type B) (Nonmuscle myosin heavy chain-B) (NMHC-B).
GN MYH10.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Ohara M., Ishiguro N., Shinaawa M.;
RT "Bos taurus nonmuscle myosin heavy chain B mRNA, complete cds."

```

Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.

123 SEQUENCE OF 204-302 FROM N.A.
124 TISSUE-Brain cortex;
125 MEDLINE-95301542; PubMed-7782316;
126 Itoh K., Adelstein R.S.;
127 Neuronal cell expression of inserted isoforms of vertebrate nonmuscle
128 myosin heavy chain II-B.
129 J. Biol. Chem. 270:14533-14540(1995).
130 -1- FUNCTION: CELLULAR MYOSIN APPEARS TO PLAY A ROLE IN CYTOKINESIS,
131 CELL SHAPE, AND SPECIALIZED FUNCTIONS SUCH AS SECRETION AND
132 CAPING (BT SIMILARITY).
133 -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY
134 CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2
135 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
136 -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPEITIVE, SHOWING
137 CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
138 CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
139 -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
140 -----
141 This SWISS-PROT entry is copyright. It is produced through a collaboration
142 between the Swiss Institute of Bioinformatics and the EMBL Outstation-
143 the European Bioinformatics Institute. There are no restrictions on its
144 use by non-profit institutions as long as its content is in no way
145 modified and this statement is not removed. Usage by and for commercial
146 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
147 or send an email to license@isb-sib.ch).
148 -----
149 EMBL: AB022023; BAA36494.1; -
150 EMBL: U15716; AAA87715.1; -
151 HSP: P08799; ILVK.
152 InterPro: IPR000048; Myosin_N.
153 InterPro: IPR002928; Myosin_tail.
154 InterPro: IPR002017; Spectrin.
155 InterPro: IPR001609; myosin_head.
156 Pfam: PF00612; IQ; 1.
157 Pfam: PF00663; myosin_head; 1.
158 Pfam: PF02276; Myosin_N; 1.
159 Pfam: PF01576; Myosin_tail; 1.
160 PRINTS: PR00193; MYOSINHEAVY.
161 ProDom: PD000355; myosin_head; 1.
162 SMART: SM00015; IQ; 1.
163 SMART: SM00242; MYSC; 1.
164 PROSITE: PS50096; IQ; 1.
165 Myosin; ATP-binding; Calmodulin-binding; Actin-binding;
166 Coiled coil; Alkylation; Multigene family.
167 KW Coiled coil; MYOSIN HEAD-LIKE.
168 FT DOMAIN 1 785
169 FT 786 815
170 FT 845 1976 COILED COIL (POTENTIAL).
171 FT NP_BIND 178 185
172 FT MOD_RES 701 701
173 FT MOD_RES 711 711 ALKYLATION (SH-1) (POTENTIAL).
174 SEQUENCE 1976 AA; 229097 MW; 614435451C0F790 CRC64;

Query Match 15.7%; Score 169; DB 1; Length 1976;
Best Local Similarity 21.4%; Pred. No. 0.027;
Matches 66; Conservative 58; Mismatches 87; Indels 98; Gaps 10;

4 INKLFDAEENV-----LDAEF--LNKELDSVKAQLSOK 38
1078 IDELKIVAKKEELQALRGDEFTLHKNNALKVYRELQAQINLEODEPSEKASNNKA 1137

39 DREKRSQAIIIDTLRDTLEERNATVESLQNALNKAEMLCSTLKK-----QMKFL 87
1138 EKQKRDLSEELAKETLELDTPTAAQDELRTKREOEVAELKALDEETKSHQAQIDOM 1197

88 EQRQ---DETKARREFAHLK-----CKMKTMOIELLLOSQR--- 122
1198 RQRIATLEELSEQLQAKKFKKLANLEKNKQGLPTDNKELACEYKAVLDQVAKSESHKRRKL 1257

123 SEVENIRDMGVQSAVEQLAVYCVSLKREYEN-----LKEARKATG----- 164
1258 DAQVQELHAKVSEGRDLRLVLEAKKANQLQNELNVSTLLEBAEKKGKIFAKDAAGLESL 1317

165 -----ELADLKKDDVSSRSKRLTNTLDELQALRLSRAQDKQASQDEI 209
1318 QDTQELLQEEFTROKLNLSRIQ-LEERSSLOEQOEEEBEAR--RSLKQALQALQAL 1373

210 TSLRKSD 218
1374 TDTKKKYD 1382

RESULT 4
G160 MOUSE STANDARD; PRT; 1325 AA.
ID G160 MOUSE
AC P55937;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Golgin-160 (Male-enhanced antigen-2) (MEA-2).
GN GOLG3 OR MEA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CD-1; TISSUE=Testis;
RC MEDLINE=97217683; PubMed=9063644;
RA Kondo M., Sutoh S.;
RT Cloning and molecular characterization of cDNA encoding a mouse
RT male-enhanced antigen-2 (Mea-2): a putative family of the Golgi
RT autoantigen.";
RL DNA Seq. 7:71-82(1997).
CC -1- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN SPERMATOGENESIS AND/OR
CC TESTIS DEVELOPMENT. PROBABLY IDENTICAL WITH THE SEROLOGICALLY
CC DETECTABLE MALE ANTIGEN (SDM).
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS. TRANSCRIPTS CAN BE
CC FOUND IN SPERMATIDS DURING SPERMATOGENESIS. NO EXPRESSION IN
CC LEYDIG CELLS, SPERMATOGENIA, OR SPERMATOCYTES.
CC -1- SIMILARITY: HIGH, TO HUMAN GOLGIN-160.
CC CAUTION: IT IS UNCERTAIN WHETHER MET-1, MET-19 OR MET-30 IS THE
CC INITIATOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D78270; BAA19612.1; -
CC HSP: P18852; ISCG.
CC MGD: MGI:96958; GOL9a3.
CC Spermatogenesis; Developmental protein.
FT DOMAIN 201 204
FT SEQUENCE 1325 AA; 149880 MW; 3230656962C687B0 CRC64;

Query Match 15.7%; Score 168.5; DB 1; Length 1325;
Best Local Similarity 23.8%; Pred. No. 0.019;
Matches 69; Conservative 50; Mismatches 84; Indels 87; Gaps 12;

10 DLAQEEENVLDA-EFLNE-----LDSVKAQLSOKDREKR-----DSQAIIIDTLRD 54
590 ELQREASREDAIIFLQNEKIVLEVALQSAKSEEDKARRRLEEDTEETSGILLEQLR 649

55 TLERNATVESLQNALNKAEMLCSTLKKOM-----KFLQD-----RODET---KQ 96
650 DLAVKSNOVEHLOE-----TATLRKQKQKVEQFQVQKVVAVYRRDQATSKDQLINE 702

Query Match	15.6%	Score 167.5	DB 1	Length 1203
Best Local Similarity	24.6%	Pred. No. 0.02		
Matches 51	Conservative 45	Mismatches 80	Indels 31	Gaps
QY	22	EFLKNELDVVAQSLQSDREKRDSQAIIITDITRLTLEERNATVESLQNALAKKAEMLCSTK	81	
		: : :	: : :	: : :

CC OF Sendmail Email to Internet@30.910.141.

DR EMBL; M35012; AAA28713.1; -
DR PIR; A36014; A36014
DR PIR; B36014; B36014.
DR HSSP; P08799; 1MMD.
DR Flybase; FBgn0005634; zip.
DR InterPro; IPR0000048; IQ.
DR InterPro; IPR002928; Myosin_N.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM000015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
KW Myosin; Alternative splicing; Coiled coil; Actin-binding;
KW Arp-binding; Calmodulin-binding;

FT DOMAIN 1 829 MYOSIN HEAD-LIKE.
 FT DOMAIN 830 859 IQ.
 FT DOMAIN 866 2017 COILED COIL (POTENTIAL).
 FT NP_BIND 225 232 ATP.
 FT DOMAIN 250 260 25 KDA/50 KDA JUNCTION.
 FT DOMAIN 682 694 50 KDA/20 KDA JUNCTION.
 FT DOMAIN 705 727 ACTIN-BINDING.
 FT DOMAIN 742 758 REACTIVE SULFHYDRYL/ACTIN-BINDING.
 FT DOMAIN 1303 2017 LIGHT MEROMYOSIN (LMM).
 FT DOMAIN 1303 1970 ALPHA-HELICAL TAILPIECE (LMM).
 FT DOMAIN 1971 2017 GLOBULAR TAILPIECE.
 FT VARSPLIC 1 45 MISSING (IN SHORT ISOFORM).
 SQ SEQUENCE 2017 AA; 232016 MW; 73E3CB02BA8F2528 CRC64;

Query Match 15.4%; Score 166; DB 1; Length 2017;
 Best Local Similarity 25.2%; Pred. No. 0.04; Indels 38; Gaps 8;

Matches 61; Conservative 47; Mismatches 96; Indels 38; Gaps 8;

QY 1 KTIINKLFEDLAQEEENVLDAAE--FLKNEUDSVKQAQISQKREKRDQSO-----AI 48
 DB 1263 KYLKEK-----AKGTLEAENADLATELRVSNSSROENDRKRAQESQIAELQYKLA 1314
 QY 49 IDTLRDTLEER---NATVESIQNALNKAEMLCSTLKKQKFLERQDQETKQAREAHN- 103
 DB 1315 IERARSELQEKCTKLOQEAENITNOLEAEALKASAVKSNMESQLEAQQLEETRQ 1374
 QY 104 ---LKKGMKMTMEQIELLOQSESEVEEMIRDMGVGSAVQQLAVCVSLKKEYNLKEAR 160
 DB 1375 KIGLSSKLRQIESEKELQLEDEDEAKRNY--EKRLAEVTTQMOEIKKKAEDADLA 1431
 QY 161 KATGELADRLKDLVSSRSKLTU---NPELDAQELRLSAQKD---LQSAQDEITSRL 213
 DB 1432 KELEBEKKRLKNQIEALERVKELIANQNDRLDKSKKQISQELSDATELEAQRKYLE 1491
 QY 214 KK 215
 DB 1492 KK 1493

RESULT 7
 USOL_YEAST STANDARD; PRT; 1790 AA.
 ID USOL_YEAST
 AC P25386;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Intracellular protein transport protein USOL.
 GN USOL OR INT1 OR YDL058W.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxID=4932;
 RN [11]
 RP SEQUENCE FROM N.A.
 RC STRAIN=X2180-1A;
 RX MEDLINE=911854402; PubMed=2010462;
 RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
 RA Yamasaki M.;
 RT "A cytoskeleton-related gene, usol, is required for intracellular
 RT protein transport in Saccharomyces cerevisiae.";
 RL J. Cell Biol. 113:245-260(1991).
 RN [12]
 RP SEQUENCE OF 782-1790 FROM N.A.
 RA Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
 RA Kendrick K.E.;
 RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
 RN [13]
 RP SEQUENCE OF 1-8 FROM N.A.
 RA Bai Y., Symington L.S.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
 CC COMPLEX.

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
 CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
 CC ER AND THE GOLGI COMPLEX.
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED
 CC OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
 CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
 CC -1- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X54378; CA38253.1; -;
 DR EMBL: L03188; AAB00143.1; -;
 DR EMBL: U53668; AAB66659.1; -;
 DR PIR: A38455; A38455.
 DR HSSP: P80220; 1DIP.
 DR SCD: S0002216; USOL.
 DR InterPro: IPR002017; Spectrin.
 DR Transprot: Protein transport; Golgi stack; Cytoskeleton; coiled coil.
 KW DOMAIN 1 724 GLOBULAR HEAD.
 FT DOMAIN 725 1790 COILED COIL (POTENTIAL).
 FT DOMAIN 465 487 CHARGED (HYPER-HYDROPHILIC).
 FT DOMAIN 921 1790 DISPENSABLE FOR THE PROTEIN FUNCTION.
 FT DOMAIN 1172 1786 ASF/GLU-RICH (ACIDIC).
 FT CONFLICT 847 847 G->E (IN REF. 2).
 FT CONFLICT 924 924 E->K (IN REF. 2).
 FT CONFLICT 1253 1253 V->I (IN REF. 2).
 FT CONFLICT 1319 1319 I->V (IN REF. 2).
 FT CONFLICT 1461 1461 N->S (IN REF. 2).
 FT CONFLICT 1581 1581 G->S (IN REF. 2).
 FT CONFLICT 1600 1600 I->V (IN REF. 2).
 FT CONFLICT 1661 1661 R->S (IN REF. 2).
 FT CONFLICT 1772 1772 D->DEEDDEE (IN REF. 2).
 SQ SEQUENCE 1790 AA; 206424 MW; 6CE2B216F9FD4818 CRC64;

Query Match 15.2%; Score 163; DB 1; Length 1790;
 Best Local Similarity 23.7%; Pred. No. 0.052;
 Matches 51; Conservative 57; Mismatches 77; Indels 30; Gaps 7;

QY 10 DLQAEENVLDAAEFLKNEUDSVKQAQISQKREKRDQSAIITDLRDLERNATVESIQNA 69
 DB 1424 ELKEEKNTIKTS--LQDEI-----LSYKDKITRNDEKLTLSIERDNKRD---LESIKEQ 1471
 QY 70 LNKAEMLCSTLKKQKFLERQDQETKQAREAHNLRKMKMTMEQIELLOQSESEVEEMI 129
 DB 1472 LRAQESKAKAVVERGLKLEBSSKEKEALERSKEM-----MKKLESTTESNEELKSSM 1525
 QY 130 RDMGVGSAVQQLAVCVSLKKEYNLKEARKATGELADRL---KDLVSSRSKLT--T 183
 DB 1526 ETIRKSDPEKLEQ-----SKKSAEEDIKNQHESKDSLIRINESEKQIEELKSLRIEAK 1579
 QY 184 LNTLELDAQELRLSAQKDQSAQDEITSRLKKSND 218
 DB 1580 SGSELETVKQELNNQAEKIRINAEENTVLKSKLED 1614

RESULT 8
 MYS2_DICDI STANDARD; PRT; 2116 AA.
 ID MYS2_DICDI
 AC P08799;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin II heavy chain, non muscle.
 GN MHCA.
 OS Dictyostelium discoideum (slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

ON NCBI_TaxID=44689;
 RX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87092266; Pubmed=3540939;
 RA Marlick H.M., de Lozanne A., Ieianwand L.A., Spudich J.A.;
 RT "Conserved protein domains in a myosin heavy chain gene from
 RT Dictyostelium discoideum";
 RL Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).
 RN [2]
 RP PHOSPHORYLATION SITES, AND MUTAGENESIS.
 RC STRAIN-AX2;
 RX MEDLINE=90353583; Pubmed=2387408;
 RA Lueck-Vielmeier D., Schlachter M., Grabatlin B., Wippler J.,
 RA Gerisch G.;
 RT "Replacement of threonine residues by serine and alanine in a
 RT phosphorylatable heavy chain fragment of Dictyostelium myosin II";
 RL FEBS Lett. 269:239-243(1990).
 RN [3]
 RP PHOSPHORYLATION SITES.
 RX MEDLINE=88112226; Pubmed=2828113;
 RA Wagle G., Noegel A., Scheel J., Gerisch G.;
 RT "Phosphorylation of threonine residues on cloned fragments of the
 RT Dictyostelium myosin heavy chain.";
 RL FEBS Lett. 227:71-75(1988).
 RN [4]
 RP PHOSPHORYLATION SITES.
 RX MEDLINE=95345066; Pubmed=7619795;
 RA Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,
 RA Rayment I.;
 RT "X-ray structures of the myosin motor domain of Dictyostelium
 RT discoidium complexed with MgADP, BeFx and MgADP.ALF4-";
 RL Biochemistry 34:8960-8972(1995).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.
 RX MEDLINE=95345067; Pubmed=7619796;
 RN Smith C.A., Rayment I.;
 RA "X-ray structure of the magnesium(II)-pyrophosphate complex of the
 RT truncated head of Dictyostelium discoidium myosin to 2.7-A
 RT resolution.";
 RL Biochemistry 34:8973-8981(1995).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
 RX MEDLINE=96206189; Pubmed=6611530;
 RA Smith C.A., Rayment I.;
 RT "X-ray structure of the magnesium(II)-ADP-vanadate complex of the
 RT Dictyostelium discoidium myosin motor domain to 1.9-A resolution.";
 RL Biochemistry 35:5404-5417(1996).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.
 RX MEDLINE=97452580; Pubmed=9305951;
 RA Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;
 RT "X-ray structures of the MgADP, MgATPgammaS, and MgAMPNP complexes
 RT of the Dictyostelium discoidium myosin motor domain.";
 RL Biochemistry 36:11619-11628(1997).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
 RX MEDLINE=98070605; Pubmed=9405148;
 RA Bauer C.B., Kuhlman P.A., Baasch C.R., Rayment I.;
 RT "X-ray crystal structure and solution fluorescence characterization
 RT of Mg₂(3')-O-(N-methylanthraniloyl) nucleotides bound to the
 RT Dictyostelium discoidium myosin motor domain.";
 RL J. Mol. Biol. 274:394-407(1997).
 RN [9]
 RP FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE
 CC ACTIVITY THAT IS ACTIVATED BY ACTIN.
 CC [10]
 RP SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
 CC INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALINE
 CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
 CC (MLC-2').
 CC [11]
 RP SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL
 CC CORTEX.
 CC [12]
 RP DOMAIN: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN BE FURTHER
 CC SPLIT INTO 2 GLOBULAR SUBSEGMENTS (S1) AND 1 ROD-SHAPED

```

CC SUBFRAGMENT (S2).
CC -I- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPERTITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -I- PPM: PHOSPHORYLATION INHIBITS THICK FILAMENT FORMATION AND REDUCES
CC THE ACTIN-ACTIVATED ATPASE ACTIVITY.
CC -I- MISCELLANEOUS: DICRYOSTELIUM MYOSIN II HAS NO K(2)EDRA ATPASE
CC ACTIVITY, PERHAPS CORRELATED WITH THE ABSENCE OF A CYS AT THE SH-1
CC POSITION (688).
CC -I- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -I- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation --
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL, M14628; AAA3227.1; -.
DR PIR, A26655; A26655.
DR PIR, S00250; S00250.
DR PDB, 1MMA: 03-DEC-97.
DR PDB, 1MMD: 17-AUG-96.
DR PDB, 1MMG: 03-DEC-97.
DR PDB, 1MMN: 03-DEC-97.
DR PDB, 1MND: 17-AUG-96.
DR PDB, 1MNE: 17-AUG-96.
DR PDB, 1VOM: 23-DEC-96.
DR PDB, 1LVK: 28-JAN-98.
DR DictyDB: DD01008; mhca.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR001609; myosin_head.
DR InterPro: IPR001612; IQ.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IQ; 1.
KW Myosin; Coiled coil; Actin-binding; ATP-binding; 3D-structure;
KW Calmodulin-binding; Methylation; Alkylation; Phosphorylation.
FT DOMAIN 1 761
FT DOMAIN 762 791
FT NP_BIND 817 2116
FT NP_BIND 179 186
FT DOMAIN 638 660
FT DOMAIN 738 752
FT MOD_RES 130 130
FT MOD_RES 678 678
FT MOD_RES 1823 1823
FT MOD_RES 1833 1833
FT MOD_RES 2029 2029
FT MOD_RES 2116 AA; 243871 MW; 2FC3770BBIIE56A1 CRC64;
SQ SEQUENCE 2116 AA; 243871 MW; 2FC3770BBIIE56A1 CRC64;

Query Match 15.1%; Score 162.5; DB 1; Length 2116;
Best Local Similarity 25.0%; Pred. No. 0.066;
Matches 67; Conservative 52; Mismatches 86; Indels 63; Gaps 11;

5 NKLFPDLAQEEENVDAEFLKNELDVSKAQALOSKD---REKRDSDAIIDLTDLLEEN 60
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
852 DLEKSLKDTESNVLDLQRLAKREKTKAMYSKDALDAEQRLREIRKEDNESELDERK 911
Oy :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 ATVESLON-----ALNKAEMLCSTLKKOMKFLERODETKQAEEEAHLKLC 106
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
912 LALELQNLQNKRSVEEKVNDLBEELQDEBQRLKNTLEKLLKKYEELEBKRRVNDG----- 965
Oy :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
107 KMKTKQELILLQSORSEVEEMI-----RDMGV-----GGSADVQLAVYCVS---- 148
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 966 QSDTISRLKIKDELQVEVELTESFSESKDKGVLEKTRVRLQSELDITVRLDSETKD 1025
QY 149 -----LKKREYENLKEARKA-TGELADRLKDLVSRSLKLTNLELDO-----AKL 193
Db 1026 KSELLRKKKLEELKQVELELAETAKLAOE--AAKKLQGEYTELNEKFNSEVTARS 1083
QY 194 ELRSQKDLQS---ADQETISLRKSD 217
Db 1084 NVEKSKKLTLESQLVAVNNLEDEKKNRD 1111

RESULT 9
RASO_THEVO
ID RASO_THEVO STANDARD; PRT; 895 AA.
AC P58302;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE DNA double-strand break repair rad50 ATPase.
GN RAD50 OR TVG0235331.
OS Thermoplasma volcanium.
OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmaceae;
OC Thermoplasma.
OX NCBI_TaxID=50339;
RN [1]
RP SEQUENCE FROM N.A.
RC SRRAIN-GSSI / DSM 4299 / JCM 9571;
RX MEDLINE=20570466; PubMed=11121031;
RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,
RA Kashiwagi-Ohya Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,
RA Ninoishi T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;
RT "Archaeal adaptation to higher temperatures revealed by genomic
RT sequence of Thermoplasma volcanium.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).
CC -1- FUNCTION: Involved in DNA double-strand break repair (DSBR). The
CC rad50/mre11 complex possesses single-strand endonuclease activity
CC and ATP-dependent double-strand-specific exonuclease activity.
CC Rad50 provides an ATP-dependent control of mre11 by unwinding
CC and/or repositioning DNA ends into the mre11 active site (By
CC similarity).
CC -1- SUBUNIT: Forms a complex with mre11 (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP000991; BAB59370.1; -
DR InterPro: IPR000439; ABC_Transport.
DR InterPro: IPR000879; Cecropin.
DR InterPro: IPR001238; RecF.
DR InterPro: IPR003395; SMC_N.
DR Pfam: PF00470; RecF. 1.
DR Pfam: PF02463; SMC_N. 1.
KW DNA repair; Hydrolase; ATP-binding; Coiled coil; Complete proteome.
FT NP_BIND 30 ATP (By similarity).
FT DOMAIN 171 737 COILED COIL (POTENTIAL).
SQ SEQUENCE 895 AA; 103223 MW; 4A47DA9287E82D3A CRC64;

Query Match 15.1%; Score 162; DB 1; Length 895;
Best Local Similarity 24.3%; Pred. NO. 0.029;
Matches 52; Conservative 52; Mismatches 80; Indels 30; Gaps 6;

```

```

Db 227 INIKENKKD---HLNEELHRLNMQLETKRYKMELEASQSRKASIEWEVVKLSIEBELK 283
QY 141 QLAAYCYSLKK-----EVENLKEARKATGELADRLK-----KLVSRSLK-- 161
Db 284 RLNNNAVVRKNEIIEINIKKDLQSLEIEIGKLSQKRYDEAHRKLELDQSRSEFLE 343
QY 182 -KTLNLELDAQLELRSAQKDLQSADEITSLRK 214
Db 344 KKKRKEDDLKRLSKLEDEDNQSAVANIENIKK 377

RESULT 10
MYS_AEOIR
ID MYS_AEOIR STANDARD; PRT; 1938 AA.
AC P24733;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, striated muscle.
DE Myosin heavy chain, striated muscle.
OS Aequipecten irradians (Bay scallop).
OC Eukaryota; Metazoa; Molusca; Bivalvia; Pteriomorpha; Pectinoidea;
OC Pectinoidea; Pectinidae; Argopecten.
OX NCBI_TaxID=31199;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adductor muscle;
RX MEDLINE=92011595; PubMed=1917970;
RA Nyitrai L., Goodwin E.B., Szent-Gyorgyi A.G.;
RT "Complete primary structure of a scallop striated muscle myosin heavy
RT chain. Sequence comparison with other heavy chain reveals regions
RT that might be critical for regulation.";
RL J. Biol. Chem. 266:18469-18476(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Adductor muscle;
RX MEDLINE=91088319; PubMed=2263488;
RA Nyitrai L., Goodwin E.B., Szent-Gyorgyi A.G.;
RT "Nucleotide sequence of full length cDNA for a scallop striated
RT muscle myosin heavy chain.";
RL Nucleic Acids Res. 18:7158-7158(1990).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 777-836.
RX MEDLINE=94173332; PubMed=8127365;
RA Xie X., Harrison D.H., Schlichting I., Sweet R.M., Kalabokis V.N.,
RA Szent-Gyorgyi A.G., Cohen C.;
RT "Structure of the regulatory domain of scallop myosin at 2.8-A
RT resolution.";
RL Nature 368:306-312(1994).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 777-836.
RX MEDLINE=96419133; PubMed=8805510;
RA Houdusse A., Cohen C.;
RT "Structure of the regulatory domain of scallop myosin at 2-A
RT resolution: implications for regulation.";
RL Structure 4:21-32(1996).
CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATOR LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC EMBL: X55714; CAA39247.1; -
DR PIR: S13557; S13557;
DR PIR: A40997; A40997;
DR PDB: 1SCM; 30-APR-94.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR002017; Spectrin.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ; 2.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IQ; 1.
KM Myosin: Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW ATP-binding; Alkylation; Calmodulin-binding; 3D-structure.
FT DOMAIN 1 777
FT DOMAIN 778 805
FT DOMAIN 836 1938
FT DOMAIN 836 1938
FT NP_BIND 176 183
FT MOD_RES 703 703
FT MOD_RES 703 703
FT HELIX 778 821
FT HELIX 822 823
FT TURN 825 833
FT TURN 834 835
SQ SEQUENCE 1938 AA; 222821 MW; A5CCEA127D1A4896 CRC64;

```

```

Query Match 15.1%; Score 162; DB 1; Length 1938;
Best Local Similarity 24.5%; Pred. No. 0.064;
Matches 73; Conservative 50; Mismatches 89; Indels 86; Gaps 12;

```

```

QY 1 KTIINKLFFD-----LAQEEENVLDLAEFLKNELDVSKAQL-----SOK 38
DB 904 KLIMQKQDFESQIKLEERLDEDAADLEGIKKMEADANAKKIDIGLENTLOKAEQ 963
QY 39 DREKRDQAIIDTL-----RDILEERN-ATVESLO-----NALKAE 74
DB 964 DKARKDNO--ISTLOGETSQODEHIIGKLNKKKALEEANKTSDSLQAEKCHLNK-- 1019
QY 75 MLCSTLKQMKFLRODDETKQAREEAHRLCKMK-----TMEQIELLLQSORSEVEEMI 129
DB 1020 -LKAKEQALDELBDNLEREKKVGDEKAKRYEQDLKSQENVEDLERVKR-ELEENV 1077
QY 130 R-----DMGVGQSAVEQLAVYCVSLKKEVENIKERKATGELADPLKLDVSSR 178
DB 1078 RRKKAETISLNSKLEDSQNLVSQLRKIKELQARIIEELELEEAERNAKAVVEQORALN 1137
QY 179 SKLTLTNLTDLQA-----KLELRSAOKDLOSA-----DOETISLRKSSD 218
DB 1138 RELEELERLDEDAAGATSAQIETLKKRREBELKTRDLEEASLOHQAISLKRKHOD 1195

```

```

RESULT 11
TANA_XENLA
ID TANA_XENLA STANDARD; PRT; 1744 AA.
AC C01550;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tanabrin.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.

```

```

OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Tadpole head;
RX MEDLINE=92398961; Pubmed=1524825;
RA Hemmati-Briavanlou A., Mann R.W., Harland R.M.;
RT "A protein expressed in the growth cones of embryonic vertebrate
neurons defines a new class of intermediate filament protein.";
RL Neuron 9:417-428(1992).
CC -1- TISSUE SPECIFICITY: GROWTH CONES OF EMBRYONIC VERTEBRATE NEURONS.
CC -1- DEVELOPMENTAL STAGE: IS EXPRESSED IN THE NEURULA AND PERSISTS
CC DURING EMBRYOGENESIS IN THE BRAIN, CRANIAL NERVES, AND SPINAL
CC CORD.
CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: M99387; AAA49966.1; -
DR PIR: JH0720; JH0720.
DR InterPro: IPR001664; IF.
DR Pfam: PF00038; filament; 2.
DR PROSITE: PS00226; IF; 1.
KW Intermediate filament; Coiled coil; Neutrone.
FT DOMAIN 1 12
FT DOMAIN 13 314
FT DOMAIN 315 1744
FT DOMAIN 8 48
FT DOMAIN 49 60
FT DOMAIN 61 156
FT DOMAIN 157 179
FT DOMAIN 180 193
FT DOMAIN 194 199
FT DOMAIN 200 314
SQ SEQUENCE 1744 AA; 199561 MW; 6502EAC9FE6CAE93 CRC64;

```

```

Query Match 14.9%; Score 160.5; DB 1; Length 1744;
Best Local Similarity 24.3%; Pred. No. 0.07;
Matches 57; Conservative 58; Mismatches 71; Indels 49; Gaps 8;

```

```

QY 12 AQEEENVLDLAEFLKNELDVSKAQLSQKREKRSQAIITDRLTLEERNATVESLQNALN 71
DB 33 ALREEN-----ELLRKETIHSLSRSSKSEKRCWKKHHEEM-KRLDALDDGH-----REMY 80
QY 72 KAEMLCSTLKQMKFLRODDETKQAREEAHRLCKMKKTWQIELLLQSOR-----TGEIADR 169
DB 81 QAENVRDSIYEIEIEFVQRCLEEKQAREDA-----KELSSKLLLEETRAQIWLKERL 135
QY 123 -----SEVEEMIRDMGVGQSAVEQLAVYCVSLKKEVENIKERKATGELADPLKLDVSSR 178
DB 136 GQLEAELELDILRHOEEKALMEE-----ELASFSQLENFRAVPAVFKVEVDVYARKLSEI 192
QY 170 LKDLVSSRSKLTNLTNLTDLQA-----KLELRSAOKDLOSA-----DOETISLRKSSD 217
DB 193 WQGAVEEYKSESVLEAGLSSEKSNLKKVLEENKQNRLLLOSIDKELVSLKMRKE 247

```

```

RESULT 12
MYSB_CAEEL
ID MYSB_CAEEL STANDARD; PRT; 1966 AA.
AC P02566;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain B (MHC B).
GN UNC-54 OR MYO-4.
OS Caenorhabditis elegans.

```


OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83273600; PubMed=6576334;
 RA Karn J., Brenner S., Barnett L.;
 RT "Protein structural domains in the *Caenorhabditis elegans* unc-54
 RT myosin heavy chain gene are not separated by introns";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).
 RN [2]
 RP SEQUENCE OF 850-1966 FROM N.A.
 RX MEDLINE=82272395; PubMed=7202124;
 RA McLachlan A.D., Karn J.;
 RT "Periodic charge distributions in the myosin rod amino acid sequence
 RT match cross-bridge spacings in muscle.";
 RL Nature 299:226-231(1982).
 RN [3]
 RP SEQUENCE OF 1876-1966 FROM N.A.
 RX MEDLINE=83232892; PubMed=6571695;
 RA Wills N., Gesteland R.F., Karn J., Barnett L., Bolten S.,
 RA Waterston R.H.;
 RT "The genes sup-7 x and sup-5 III of *C. elegans* suppress amber
 RT nonsense mutations via altered transfer RNA.";
 RL Cell 33:575-583(1983).
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (LMC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LMN) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -1- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN
 CC C.ELEGANS.
 CC -1- MISCELLANEOUS: MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY
 CC WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: J01050; AAA28124.1; -;
 DR EMBL: V01494; CAA24738.1; -;
 DR PIR: A02992; MKMW.
 DR HSSP: P08799; 1MND.
 DR InterPro: IPR004009; Myosin_N.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00063; myosin_head. 1.
 DR Pfam: PF02736; Myosin_N. 1.
 DR Pfam: PF01576; Myosin_tail. 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRODOM: PD000355; myosin_head. 1.
 DR SMART: SM00242; MYSC; 1.
 KM APP-binding: Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KM APP-binding: Methylation; Alkylation; Multigene family.
 FT DOMAIN 1 850 MYOSIN HEAD-LIKE.
 FT DOMAIN 851 1966 COILED COIL (POTENTIAL).
 FT DOMAIN 851 1164 ALPHA-HELICAL TAILPIECE (S2).
 FT DOMAIN 1165 1176 HINGE.
 FT DOMAIN 1165 1966 LIGHT MEROMYOSIN (LMN).

FT NP_BIND 177 184 ATP (BY SIMILARITY).
 FT DOMAIN 665 687 ACTIN-BINDING.
 FT DOMAIN 769 783 ACTIN-BINDING.
 FT MOD_RES 128 128 METHYLATION (TRI-) (POTENTIAL).
 FT MOD_RES 705 705 ALKYLATION (SH-1).
 FT MOD_RES 715 715 ALKYLATION (SH-2).
 FT CONFLICT 1337 1337 E -> R (IN REF. 2).
 FT CONFLICT 1880 1880 I -> L (IN REF. 2).
 FT SEQUENCE 1966 AA; 225125 MM; B66F0BB2FE27B67F CXC64;
 SQ
 Query Match 14.9%; Score 160.5; DB 1; Length 1966;
 Best Local Similarity 25.1%; Pred. No. 0.079;
 Matches 59; Conservative 41; Mismatches 78; Indels 57; Gaps 9;
 QY 3 IINKLFDLAOEENVDIAEFLKNELDSVKAQLSQKREKRDSQALITDITRDILEERNAT 62
 DB 1014 INRKLMEDLOSSEDK-----GNHQNKKVAKLEQ-----TLDDLEDSLREKRA 1056
 QY 63 VESLQNALNKAEMLCSFLKQMKFLRODET-KQAREEAHLKCKKMTQETLELLQSQ 121
 DB 1057 RADDKQKKRVE-----GELKTAQENIDSGQRHDLNNLKKSELSHVSRLDE 1109
 QY 122 RSEYEEMIRDMGYQGSAAVEDLAVYCVSLKREYENLKEARRATGELADRLKDIYSSRSKL 181
 DB 1110 QALVSKLRQRIKQDSRISL-----LEELERENRSRCK-----ADRAKSDL---QREL 1155
 QY 182 KTLNTELDQ-----AKLELRSAQKDLQSA---DOETISLRKSSD 218
 DB 1156 EELGEKLEDEGGATAAAGVYKKREAEALAKLRDLLEENAMNHNENQGLGRLKKHTD 1210
 RESULT 13
 MYHB_MOUSE STANDARD; PRT; 1972 AA.
 ID MYHB_MOUSE
 AC 008638; 008639; 062462; 064195;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin heavy chain, smooth muscle isoform (SMHMC).
 GN Myh11.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPRAIN-BALB/C; TISSUE=Uterus;
 RX MEDLINE=97242182; PubMed=9125171;
 RA Hasegawa K., Arakawa E., Oda S., Matsuda Y.;
 RT "Molecular cloning and expression of murine smooth muscle myosin heavy
 RT chains";
 RL Biochem. Biophys. Res. Commun. 232:313-316(1997).
 RN [2]
 RP SEQUENCE OF 1-368 FROM N.A.
 RC TISSUE=Uterus;
 RX MEDLINE=95008063; PubMed=7923625;
 RA Milano J.M., Cserjesi P., Liqon K.L., Perissamy M., Olson E.N.;
 RT "Smooth muscle myosin heavy chain exclusively marks the smooth muscle
 RT lineage during mouse embryogenesis";
 RL Circ. Res. 75:803-812(1994).
 RN [3]
 RP SEQUENCE OF 1-126 FROM N.A.
 RC TISSUE=Uterus;
 RX MEDLINE=96172919; PubMed=8593698;
 RA Suzuki T., Kim H.S., Kurabayashi M., Hamada H., Fujii H., Aikawa M.,
 RA Watanabe M., Watanabe N., Sakomura Y., Yazaki Y., Nagai R.;
 RT "Preferential differentiation of p19 mouse embryonal carcinoma cells
 RT into smooth muscle cells. Use of retinoic acid and antisense against
 RT the central nervous system-specific POU transcription factor Brn-2.";
 RL Circ. Res. 78:395-404(1996).
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2

```

CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 (SHOWN HERE) AND 2: ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PPM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC AKTYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHADED
CC SUBFRAGMENT (S2).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D85923; BAA19690.1; -
DR EMBL; D85924; BAA19691.1; -
DR EMBL; L25860; AAA67552.1; -
DR EMBL; S81516; AAB36168.1; -
DR HSP; P10587; IBL1.
DR MGD; MGI:102643; MYH11.
DR InterPro: IPR000048; IO.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR001609; myosin_head.
DR Pfam; PF00612; IO; 1.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRODOM; PD000355; myosin_head; 1.
DR SMART; SM00015; IO; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IO; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
KW Multigene family; Alternative splicing.
FT DOMAIN 1 785 MYOSIN HEAD-LIKE.
FT DOMAIN 786 815 IO.
FT DOMAIN 844 1934 COILED COIL (POTENTIAL).
FT DOMAIN 1935 1972 CARBOXYL-TERMINAL.
FT NP_BIND 178 185 ATP (POTENTIAL).
FT DOMAIN 661 683 ACTIN-BINDING (BY SIMILARITY).
FT DOMAIN 762 776 ACTIN-BINDING (BY SIMILARITY).
FT MOD_RES 129 129 METHYLATION (TR1-) (POTENTIAL).
FT MOD_RES 701 701 ALKYLATION (SH-1) (POTENTIAL).
FT MOD_RES 711 711 ALKYLATION (SH-2) (POTENTIAL).
FT VARSPPLIC 1930 1972 RGNASFPVSRAGRGVIENTDSEEMARDFNGTKA
SE -> GPPQETSQ (IN ISOFORM 2).
FT CONFLICT 126 126 N -> D (IN REF. 3).
FT CONFLICT 161 161 A -> V (IN REF. 2).
FT CONFLICT 189 189 Q -> K (IN REF. 2).
SQ SEQUENCE 1972 AA; 227026 MW; A1398B3F5B1F15A CRC64;

```

Query Match 14.98; Score 160; DB 1; Length 1972;
 Best Local Similarity 22.38; Pred. No. 0.084;
 Matches 53; Conservative 53; Mismatches 91; Indels 48; Gaps 7;

OY 10 DLAAEEVVDLAEFLKNELDSVKRAQLSK-----DREKRDQAIIITDLRLTEE 58
 Db 1688 DLMLQEDLAAAEKRAKQADLEKEELAEELASSLSGRNTLQDEKRRRLRLAQLLEEL 1747

```

OY 59 RNATVESIQNALNKAEMLCSTLKKOMKFLERODETKOAREEARH----- 103
Db 1748 EOGNWEAMSDVRKATLQAEQLSNELATERSTAQKNESARQOLEROKKELRSKLQVEGA 1807
OY 104 LKCMK-TMEDIELLOSREVEEMIRDMKVGSAVQOLAVYVSLKKEVENLKEARKA 162
Db 1808 VKARLKSTVAALAEKAIQLEQVQVQEARER---QAAKRSLKQKDKKLEVLQVEDERK- 1863
OY 163 TGEIADRLLKQDLVSRSKSLTKTLNTELOAKLE-----LRSADKQDSADQ-----E 208
Db 1864 ---MAEQYKDAEKGNGKTKVKQLKQLEBAEBSOCIANNRKRLQREIDATEESNEAGRE 1920
OY 209 ITSLLRK 215
Db 1921 VNALKSK 1927

```

```

RESULT 14
KINH_HUMAN STANDARD; PRT; 963 AA.
AC P33176;
DT 01-OCT-1993 (Rel. 27, Created)
DT 16-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Kinesin heavy chain (ubiquitous kinesin heavy chain) (UKHC).
GN KIF5B OR KNS1 OR KNS.
OS Homo sapiens (Human).
OC Eukaryota; Euteleostomi; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=92296683; PubMed=1607388;
RA Navone F., Nicolas J., Hom-Booher N., Sparks L., Bernstein H.D.,
RA McCallfrey G., Vale R.D.;
RT Cloning and expression of a human kinesin heavy chain gene:
RT Interaction of the COOH-terminal domain with cytoplasmic microtubules
RT in transfected CV-1 cells."
RL J. Cell Biol. 117:1263-1275(1992).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 1-349.
RX MEDLINE=96195066; PubMed=8606779;
RA Kull F.J., Sablin E.P., Lau R., Fletcher R.J., Vale R.D.;
RT "Crystal structure of the kinesin motor domain reveals a structural
RT similarity to myosin."
RL Nature 380:550-555(1996).
RN [3]
RP TISSUE SPECIFICITY.
RX MEDLINE=94242426; PubMed=7514426;
RA Nicolas J., Navone F., Hom-Booher N., Vale R.D.;
RT Cloning and localization of a conventional kinesin motor expressed
RT exclusively in neurons."
RL Neuron 12:1059-1072(1994).
CC -1- FUNCTION: KINESIN IS A MICROTUBULE-ASSOCIATED FORCE-PRODUCING
CC PROTEIN THAT MAY PLAY A ROLE IN ORGANELLE TRANSPORT.
CC -1- SUBUNIT: OLIGOMER COMPOSED OF TWO HEAVY CHAINS AND TWO LIGHT
CC CHAINS.
CC -1- SUBCELLULAR LOCATION: UNIFORMLY DISTRIBUTED BETWEEN THE CELL BODY
CC AND THE PROCESSES IN THE NEURONS.
CC -1- TISSUE SPECIFICITY: FOUND IN NEWBORN AND ADULT BRAIN, LIVER,
CC KIDNEY, SPLEEN, HEART, LUNG AND SCIATIC NERVE.
CC -1- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS: A LARGE GLOBULAR N-
CC TERMINAL DOMAIN WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY OF
CC KINESIN (IT HYDROLYZES ATP AND BINDS MICROTUBULE), A CENTRAL
CC ALPHA-HELICAL COILED COIL DOMAIN THAT MEDIATES THE HEAVY CHAIN
CC DIMERIZATION; AND A SMALL GLOBULAR C-TERMINAL DOMAIN WHICH
CC INTERACTS WITH OTHER PROTEINS (SUCH AS THE KINESIN LIGHT CHAINS),
CC VESICLES AND MEMBRANOUS ORGANELLES.
CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```


FT DOMAIN 1 785 MYOSIN HEAD-LIKE.
 FT DOMAIN 786 815 IO.
 FT DOMAIN 844 1934 COILED COIL (POTENTIAL).
 FT DOMAIN 1935 1972 CARBOXYL-TERMINAL.
 FT NR_BIND 178 185 ATP (POTENTIAL).
 FT DOMAIN 661 683 ACTIN-BINDING (BY SIMILARITY).
 FT DOMAIN 762 776 ACTIN-BINDING (BY SIMILARITY).
 FT MOD_RES 129 129 METHYLATION (SH-1) (POTENTIAL).
 FT MOD_RES 701 701 ALKYLATION (SH-2) (POTENTIAL).
 FT MOD_RES 711 711 ALKYLATION (SH-2) (POTENTIAL).
 FT CONFLICT 887 889 EEL -> NSE (IN REF. 3).
 FT CONFLICT 1263 1266 ELOS -> TISE (IN REF. 2).
 FT CONFLICT 1558 1558 T -> S (IN REF. 3).
 FT CONFLICT 1610 1611 KO -> NE (IN REF. 3).
 FT CONFLICT 1786 1786 A -> S (IN REF. 4).
 FT CONFLICT 1958 1958 T -> L (IN REF. 3).
 SQ SEQUENCE 1972 AA; 227338 MW; 67655BB2ACE1277 CRC64;

Query Match 14.8%; Score 159; DB 1; Length 1972;
 Best Local Similarity 23.9%; Pred. No. 0.095;
 Matches 59; Conservative 52; Mismatches 88; Indels 48; Gaps 9;

QY 10 DLAGEEENVLDALFLKNELDVKAQLSQK-----DREKDSQAIDTLRDTLEE 58
 DB 1688 DLMOJEDDLAAERARQADLEKELEBELASSLSGRNALODEKRRLERLAEIAQLEEELEE 1747
 QY 59 RNATVESLONALNK-----AEMLC-----STLKKOMFLEORODETKOAREEHAHL--- 104
 DB 1748 EOGNMEASDVYRKATQOAEOLSNELATERSTAOKNESARQOLEKONELSKLHEMGA 1807
 QY 105 -KCKMK-TMEOIELLSQSEVEEMIRDMGVGSAVEQLAVYCVSLKKEYENKEARKA 162
 DB 1808 VKSKFKSTIALLEAKIQLQEVQGEAREK---QAATKSLKQKKKLKEILLQYDEDERK- 1863
 QY 163 TGEIADLRKDKLVSSRSKLTLTNLTLDQAKLE-----LRSQKDLQASDQ-----E 208
 DB 1864 ---MAEOYKEQAEKGNKAVKOLKROLEEAEESORINANRRKQLRELDATESNEAMGRE 1920
 QY 209 ITSLSRK 215
 DB 1921 VNALSK 1927
 RESULT 16
 MYHB_RABIT
 ID MYHB_RABIT STANDARD; PRT; 1972 AA.
 AC P35748;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin heavy chain, smooth muscle isoform (SMHMC).
 GN MYH1.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=92073350; Pubmed=1961735;
 RA Babilj P., Kelly C., Perlasamy M.;
 RT "Characterization of a mammalian smooth muscle myosin heavy-chain
 RT gene: complete nucleotide and protein coding sequence and analysis of
 RT the 5' end of the gene";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:10676-10680(1991).
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (LHC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MHC-2).
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.

CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; M77812; AAA31395.1; -;
 DR PIR; A41604; A41604.
 DR HSSP; P08799; IMMD.
 DR InterPro; IPR000048; IO.
 DR InterPro; IPR004009; Myosin_N.
 DR InterPro; IPR002928; Myosin_Tail.
 DR InterPro; IPR002017; Spectrin.
 DR Pfam; PF00612; IQ; 1.
 DR Pfam; PF00663; myosin_head; 1.
 DR Pfam; PR02736; Myosin_N; 1.
 DR Pfam; PR01576; Myosin_Tail; 1.
 DR PRINTS; PR00193; MYOSINHEAVY.
 DR ProDom; PD000355; myosin_head; 1.
 DR SMART; SM00015; IQ; 1.
 DR SMART; SM00242; MYSC; 1.
 DR PROSITE; PSS0096; IQ; 1.
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
 KW Multigene family.
 FT DOMAIN 1 784 MYOSIN HEAD-LIKE.
 FT DOMAIN 785 807 IO.
 FT DOMAIN 844 1934 COILED COIL (POTENTIAL).
 FT DOMAIN 1935 1972 CARBOXYL-TERMINAL.
 FT NR_BIND 178 185 ATP (POTENTIAL).
 FT DOMAIN 661 683 ACTIN-BINDING (BY SIMILARITY).
 FT DOMAIN 763 777 ACTIN-BINDING (BY SIMILARITY).
 FT MOD_RES 129 129 METHYLATION (SH-1) (POTENTIAL).
 FT MOD_RES 701 701 ALKYLATION (SH-1) (POTENTIAL).
 FT MOD_RES 711 711 ALKYLATION (SH-2) (POTENTIAL).
 SQ SEQUENCE 1972 AA; 227318 MW; 2061A22428BDA4C CRC64;

Query Match 14.8%; Score 159; DB 1; Length 1972;
 Best Local Similarity 22.7%; Pred. No. 0.095;
 Matches 56; Conservative 52; Mismatches 91; Indels 48; Gaps 7;

QY 10 DLAGEEENVLDALFLKNELDVKAQLSQK-----DREKDSQAIDTLRDTLEE 58
 DB 1688 DLMOJEDDLAAERARQADLEKELEBELASSLSGRNALODEKRRLERLAEIAQLEEELEE 1747
 QY 59 RNATVESLONALNK-----AEMLCSTLKKOMFLEORODETKOAREEHAHLKCKMKTVE-- 112
 DB 1748 EOGNMEASDVYRKATQOAEOLSNELATERSTAOKNESARQOLEKONELSKLQEMGA 1807
 QY 113 -----QIELLSQSEVEEMIRDMGVGSAVEQLAVYCVSLKKEYENKEARKA 162
 DB 1808 VKSKFKSTIALLEAKIQLQEVQGEAREK---QAAKALKQDKKLKEMILQVEDERK- 1863
 QY 163 TGEIADLRKDKLVSSRSKLTLTNLTLDQAKLE-----LRSQKDLQASDQ-----E 208
 DB 1864 ---MAEOYKEQAEKGNKAVKOLKROLEEAEESORINANRRKQLRELDATESNEAMGRE 1920
 QY 209 ITSLSRK 215
 DB 1921 VNALSK 1927

```

RESULT 17
RA50_ARCFU STANDARD; PRT; 886 AA.
ID RA50_ARCFU
AC 029230;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA double-strand break repair rad50 ATPase.
GN RAD50 OR AF1032.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Kleen H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Richardson K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
RA Fleischmann D.L., Kertavage A.R., Graham D.E., Kyrpides N.C.,
RA Kirschner E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Goeyne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
RA Sedow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -!- FUNCTION: Involved in DNA double-strand break repair (DSBR). The
CC rad50/mre11 complex possesses single-strand endonuclease activity
CC and ATP-dependent double-strand-specific exonuclease activity.
CC Rad50 provides an ATP-dependent control of mre11 by unwinding
CC and/or repositioning DNA ends into the mre11 active site (by
CC similarity).
CC -!- SUBUNIT: Forms a complex with mre11 (by similarity).
CC -!- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE001032; MAB90211.1; -.
DR TIGR: AF1032;
DR InterPro: IPR003439; ABC_transport.
DR InterPro: IPR001238; RecF.
DR InterPro: IPR002017; Spectrin.
DR Pfam: PF00470; RecF. 1.
KM DNA repair; Hydrolase; ATP-binding; Coiled coil; Complete proteome.
FT NP_BIND 31 728 ATP (BY SIMILARITY).
FT DOMAIN 148 728 COILED COIL (POTENTIAL).
SQ SEQUENCE 886 AA; 103633 MW; D35641D499AAB58 CRC64;

```

Query Match 14.7%; Score 157.5; DB 1; Length 886;
 Best Local Similarity 22.3%; Pred. No. 0.051;
 Matches 56; Conservative 71; Mismatches 83; Indels 41; Gaps 9;

```

QY 1 KTIINKLFFDLAEEVNLDAEFLKNEIDSVKAO-----LSQKDRK-KRDSQAI 49
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 207 ESLKEKSEEEVNRNLSRLKEEHSKRLSESRKQESSVLQEVRLLEKRLLEQLKEV 266
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 50 DTLADTLERRAIVESLONALNK---ADMCLSTLKKQMKFLQNR-----QDET 95
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 267 ERIED-LEKKAKEVLEKPKAERYSLLEKLSEINQALRDVEKRECDLTRRAGIQOLK 325
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

QY 96 QAREARHLCKMKTMEQIE-----LLLOSREVEEMTRDQSGSAVEQLAVY 145
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 326 KAEDNSKLEETTRITELERELERFEKSHRLTTLKPKMDR---OGI-KAKLEENLT 381
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 146 CVSLKEEYENLEKARKATGELADRLKDLVSSRSKLTMTLWELDOAKLELSAQDKQSA 205
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 382 PDKKEKMYDLSLKKAKEKEKTEKTKK-LIAKKSLSLTKRGAQLKKAIVELKSNRTPVC 440
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY 206 DQETLSLRKS 216
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 441 GREIDEEHRKN 451
   :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

```

RESULT 18
MYSN_ACACA STANDARD; PRT; 1509 AA.
ID MYSN_ACACA
AC P05659;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin II heavy chain, non muscle.
OS Acanthamoeba castellanii (Amoeba).
OC Eukaryota; Acanthamoebidae; Acanthamoeba.
OX NCBI_TaxID=5755;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87308395; PubMed=3040773;
RA Hammer J.A., III, Bowers B., Paterson B.M., Korn E.D.;
RT "Complete nucleotide sequence and deduced polypeptide sequence of a
RT nonmuscle myosin heavy chain gene from Acanthamoeba: evidence of a
RT hinge in the rodlike tail."
RL J. Cell Biol. 105:913-925(1987).
CC -!- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO F-ACTIN & HAS APPARE
CC ACTIVITY THAT IS ACTIVATED BY F-ACTIN.
CC -!- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
CC INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI
CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
CC (MLC-2).
CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC OF ALPHA-HELICAL STRUCTURES. THIS REGION IS
CC INTERRUPTED BY A HINGE AND JOINED BY A NONHELICAL TAILPIECE WHERE
CC THE REGULATORY PHOSPHORYLATION SITES RESIDE.
CC -!- MISCELLANEOUS: THE HINGE REGION MAY PLAY A KEY ROLE IN MEDIATING
CC THE EFFECT OF HEAVY CHAIN PHOSPHORYLATION ON ENZYMIC ACTIVITY.
CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 IQ DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y00624; CA68663.1; -.
DR PIR: A27224; A27224.
DR HSSP: P08799; IMND.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ; 2.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MISC; 1.
DR PROSITE: PS50096; IQ; 1.
KM Myosin; Coiled coil; Actin-binding; ATP-binding; Calmodulin-binding;
KW Methylation; Alkylation; Phosphorylation; Multigene family.

```

```
FT DOMAIN 1 789 MYOSIN HEAD-LIKE.
FT DOMAIN 790 819 IQ.
FT DOMAIN 848 1509 COILED COIL (POTENTIAL).
FT DOMAIN 848 1226 ALPHA-HELICAL TAILPIECE (S2).
FT DOMAIN 1227 1252 HINGE.
FT DOMAIN 1253 1509 LIGHT MEROMYOSIN (LMM).
FT DOMAIN 1253 1482 ALPHA-HELICAL TAILPIECE (LMM).
FT DOMAIN 1483 1509 NONHELICAL TAILPIECE.
FT NP_BIND 182 189 ATP.
FT DOMAIN 660 682 ACTIN-BINDING.
FT DOMAIN 766 780 ACTIN-BINDING.
FT MOD_RES 133 133 METHYLATION (TR1-) (POTENTIAL).
FT MOD_RES 700 700 AKYLATION (SH-1) (POTENTIAL).
FT MOD_RES 1489 1489 PHOSPHORYLATION.
FT MOD_RES 1494 1494 PHOSPHORYLATION.
FT MOD_RES 1499 1499 PHOSPHORYLATION.
SO SEQUENCE 1509 AA; 171201 MW; 2CE49BE51173D17E CRC64;

Query Match 14.6%; Score 156.5; DB 1; Length 1509;
Best Local Similarity 20.9%; Pred. No. 0.099;
Matches 59; Conservative 59; Mismatches 91; Indels 73; Gaps 7;

QY 10 DLADAEENVLD-----AEFLKNEIDSV-----KAQLS 36
D 1017 DALDAENISFELSKLNTKTERGADVNEIDVATATKLEKTKSLSEELATRAOLE 1076
QY 37 QKDEKRSQALIDTLRDLERNATVESLONALNKAEMLCSTLKKOMKFLERODETKQ 96
D 1077 EEKSGKEAASKAKOLGOOLEANSEVDLSKLSAASKSLKTADONRDLDEOLEDERT 1136
QY 97 AR---EEAHLKCKMKMTMEQIELLQSO-----RSEVEEMRDGVGQ----- 136
D 1137 VRANVDKOKKALKLEKTELEQVYALDQKNAAQAQKTLKQVDETRRLLEEAASAR 1196
QY 137 -----SAVEQLAYCVSLKEYEENLEKARKATGELADRLKKDLV-----SSRSKLT 183
D 1197 LEKERKNALDEVAQLTADLDERDSGAQORRKLNRISLESELENAPKRTGASSEYVKR 1256
QY 184 LTELDOAKLELRSQA-----KDLQSADEITSLRKSDD 218
D 1257 LEGELERLEELTRAQEARAAAEKNLKANLELELROEAD 1298

RESULT 19
VDP_HUMAN STANDARD; PRT; 962 AA.
ID VDP_HUMAN
AC 060763;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE General vesicular transport factor p115 (Transcytosis associated
  protein) (TAP) (Vesicle docking protein).
GN VDP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A. AND PHOSPHORYLATION SITE SER-942.
RX MEDLINE-98148093; Pubmed-9478999;
RA Sonoda M., Misumi Y., Yano A., Takami N., Ikehara Y.;
RT "Phosphorylation of the vesicle docking protein p115 regulates its
  association with the Golgi membrane."
RT J. Biol. Chem. 273:5385-5388(1998).
RL -1- FUNCTION: GENERAL VESICULAR TRANSPORT FACTOR REQUIRED FOR
  TRANSCYTOTIC FUSION AND/OR SUBSEQUENT BINDING OF THE VESICLES TO
  THE TARGET MEMBRANE. MAY WELL ACT AS A VESICULAR ANCHOR BY
  INTERACTING WITH THE TARGET MEMBRANE AND HOLDING THE VESICULAR
  AND TARGET MEMBRANES IN PROXIMITY (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: PERIPHERAL MEMBRANE PROTEIN WHICH RECYCLES
  BETWEEN THE CYTOSOL AND THE GOLGI APPARATUS DURING INTERPHASE.
```

```
CC -1- DOMAIN: COMPOSED OF A GLOBULAR HEAD, AN ELONGATED TAIL (COILED-
  COIL) AND A HIGHLY ACIDIC C-TERMINAL DOMAIN.
CC -1- PTM: PHOSPHORYLATED IN A CELL CYCLE-SPECIFIC MANNER:
  PHOSPHORYLATED IN INTERPHASE BUT NOT IN MITOTIC CELLS.
CC DEPHOSPHORYLATED PROTEIN ASSOCIATES WITH THE GOLGI MEMBRANE:
  PHOSPHORYLATED PROMOTES DISSOCIATION.
CC -1- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC EMBL; D86326; BAA25300.1; -.
DR MIM; 603344; -.
DR InterPro; IPR00225; Armadillo.
DR PROSITE; PS50176; ARM_REPEAT; UNKNOWN_1.
KW Transport; Protein transport; Golgi stack; Membrane; Coiled coil;
  Phosphorylation.
FT DOMAIN 1 637 GLOBULAR HEAD.
FT DOMAIN 638 930 COILED COIL (POTENTIAL).
FT DOMAIN 935 962 ASP/GLU-RICH (ACIDIC).
FT MOD_RES 942 942 PHOSPHORYLATION.
FT MUTAGEN 942 942 S->A: LOSS OF PHOSPHORYLATION.
SO SEQUENCE 962 AA; 107906 MW; 2E748FC1BC2B942 CRC64;

Query Match 14.4%; Score 155; DB 1; Length 962;
Best Local Similarity 19.8%; Pred. No. 0.076;
Matches 64; Conservative 57; Mismatches 93; Indels 110; Gaps 6;

QY 3 IINKLFPDLADAEENVLDLAEFLKNEIDSVKQAOLSKDREKDSQAIIIDTLDEERNAT 62
D 608 IFDEHFTLVLELGVITKAYIKSSEEDKKEEVKKTLEQHDN--IYHYKNMIREDDIQ 665
QY 63 VESLONALN-----KAEMLCSTLKKOMKFLERODETKQAR----- 98
D 666 LEELRQOVSTLKCQNEOLQAVTQOVSOIQHKQYINLTKIQLGKDNHOGSYSEGAOMN 725
QY 99 ---EEAHLKCKMKMTMEQIELLQSOSEVEEMIRDMVGOSA----- 138
D 726 GIOPEEIGRLREIELELKRNOELLQSLTEKDSMIENKKSOTSQTNBOSSAIVARDE 785
QY 139 -VEQLAYCVSLK----- 150
D 786 QVAELKDELATLKSQNSQVEITKLOTKEQLLQKTEAFKSVVOGETTITATKTTD 845
QY 151 -----KEYENLEKARKATGELADRLKKDLVSSRSKLTLMNTELDQAKLELRSQKD 201
D 846 VEGRLSALLOQTEKELKNKRIKLSERRAIKQDLOSSNSTIAILQTEKDKELEITDSKE 905
QY 202 -----LQSADEITSLRKSDD 218
D 906 QDDLVLVLLADQDKITLSLKNLKD 929

RESULT 20
MYS_A_DROME STANDARD; PRT; 2411 AA.
ID MYS_A_DROME
AC P05661;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, muscle.
GN MHC.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
```

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89384556; PubMed=2506434;
 RA George E.L., Ober M.B., Emerson C.P. Jr.;
 RT "Functional domains of the Drosophila melanogaster muscle myosin
 heavy-chain gene are encoded by alternatively spliced exons."
 RL Mol. Cell. Biol. 9:2957-2974(1989).
 RN [2]
 RP SEQUENCE OF 1-312 FROM N.A.
 RX MEDLINE=87280141; PubMed=3038896;
 RA Wassenberg D.R. II, Kronert W.A., O'Donnell P.T., Bernstein S.I.;
 RT "Analysis of the 5' end of the Drosophila muscle myosin heavy chain
 gene. Alternatively spliced transcripts initiate at a single site and
 intron locations are conserved compared to myosin genes of other
 organisms."
 RT J. Biol. Chem. 262:10741-10747(1987).
 RN [3]
 RP SEQUENCE OF 486-881 FROM N.A.
 RC STRAIN=CANTON-S; TISSUE=Embryonic muscle;
 RX MEDLINE=91330870; PubMed=1907912;
 RA Kronert W.A., Edwards K.A., Roobe E.S., Wells L., Bernstein S.I.;
 RT "Muscle-specific accumulation of Drosophila myosin heavy chains: a
 splicing mutation in an alternative exon results in an isoform
 substitution."
 RT EMBO J. 10:2479-2488(1991).
 RL
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -1- ALTERNATIVE PRODUCTS: AS DROSOPHILA HAS A SINGLE MUSCLE MHC GENE,
 MHC ISOFORMS MUST RELY ON ALTERNATIVE RNA SPLICING THAT LEADS TO
 DIFFERENCES IN THE C-TERMINUS OF THE VARIOUS MHC PROTEINS.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
 SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 SUBFRAGMENT (S2).
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: M61229; AAA28686.1; ALT_SEQ.
 DR EMBL: M61229; AAA28687.1; ALT_SEQ.
 DR EMBL: J02788; AAA28706.1; ALT_SEQ.
 DR EMBL: J02788; AAA28707.1; ALT_SEQ.
 DR EMBL: X60196; CAA42752.1; ALT_SEQ.
 DR EMBL: X60196; CAA42753.1; ALT_SEQ.
 DR EMBL: X60196; CAA42754.1; ALT_SEQ.
 DR PIR: A28492; A28492.
 DR PIR: A32491; A32491.
 DR PIR: B32491; B32491.
 DR HSSP: P08799; 1MND.
 DR FlyBase: Fggn0002741; Mhc.
 DR InterPro: IPR000048; IO.
 DR InterPro: IPR004009; Myosin_N.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IO; 2.
 DR Pfam: PF00063; myosin_head; 7.
 DR Pfam: PF02736; Myosin_N; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRODOM: PD000355; myosin_head; 5.
 DR SMART: SM00015; IO; 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS50096; IO; 1.

KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KM ATP-binding; Calmodulin-binding; Alternative splicing;
 KW Multigene family.
 FT DOMAIN 1
 FT DOMAIN 1205 1232
 FT DOMAIN 1233 2376
 FT NP_BIND 227 234
 FT VARSPIC 2385 2385
 FT VARSPIC 2386 2411
 FT CONFLICT 443 444
 SQ SEQUENCE 2411 AA: 276434 MW: 804FEC1BB8D510A9 CRC64;
 Query Match 14.4%; Score 154.5; DB 1; Length 2411;
 Best Local Similarity 22.8%; Pred. No. 0.21; Indels 53; Gaps 9;
 Matches 59; Conservative 56; Mismatches 53; Gaps 9;
 QY 10 DLAOEENVD--AEFLKNEIDSVKQQLSQDKREKRDQAIIIDRLPTLEERNATVESIQ 67
 Db 2113 DDAREQGISERRANMLQNELESRTLLLEGADRGRAQEGELADAHQELNEVSQAQNASIS 2172
 QY 68 NALNKAMLCSTLKQKQFL---EQRDETKAREBAHRLKCKMTMEDIELLOSQRS 123
 Db 2173 AAKRKLESELQTLHSDLELLNEAKNSEKAKRAMVDAARLADELRAEDDHAQTOEKLK 2232
 QY 124 EYEMIRDMGV-----GQSAVEQLAVYCVSLKKEYE-----NLEKAR 160
 Db 2233 ALDQIKELQVRLDEAFANLKGGRKATQKLEQRYLELLEDEQRRNADQKNLKSE 2292
 QY 161 KATGELA----DLK---KDLVSS-RSKLKTLMLELDAQ-----DELRSQKDLQ 203
 Db 2293 RYVKELSEFQSEEDRKRNHERMQDLYDKLQKIKYKRGIEBAEELALNLAKPKQAQELE 2352
 QY 204 S-----ADQETSLKK 215
 Db 2353 EAERADLAQALSKPRAK 2371
 RESULT 21
 ID EP15_HUMAN STANDARD; PRT; 896 AA.
 AC P42566;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Epidermal growth factor receptor substrate 15 (Protein EP15) (AF-1p
 protein).
 GN EP15 OR AF1P.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=melanoma;
 RX MEDLINE=94239734; PubMed=8183552;
 RA Wong W.T., Kraus M.H., Carlomagno F., Zelano A., Druck T.,
 RA Croce C.M., Huebner K., di Fiore P.P.;
 RT "The human eps15 gene, encoding a tyrosine kinase substrate, is
 RT conserved in evolution and maps to 1p31-p32."
 RL Oncogene 9:1591-1597(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94181254; PubMed=8134107;
 RA Bernard O.A., Mauchault M., Mecucci C., van den Berghe H.,
 RA Berger R.;
 RT "A novel gene, AF-1p, fused to HRX in t(1;11)(p32;q23), is not
 RT related to AF-4, AF-9 nor ENL."
 RL Oncogene 9:1039-1045(1994).
 CC -1- FUNCTION: INVOLVED IN CELL GROWTH REGULATION. MAY PLAY A ROLE IN
 CC SIGNAL TRANSDUCTION AND MITOGENICITY.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED.

```

CC -1- P-TM: PHOSPHORYLATED ON TYROSINE BY EGFR.
CC -1- DISEASE: INVOLVED IN A T(1;11)(P32;Q23) CHROMOSOMAL TRANSLOCATION
CC IN ACUTE LEUKEMIAS CAUSING FUSION TO THE TRITHORAX (MLL OR HRX)
CC GENE PRODUCT WHICH CONTAINS DNA-BINDING MOTIFS RESULTING IN A
CC ROGUE ACTIVATOR PROTEIN.
CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U07707; AAA52101.1; -.
DR EMBL: Z29064; CAA82305.1; -.
DR HSSP: P02633; 1BOD.
DR MIM: 600051; -.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR000261; EF-hand.
DR InterPro: IPR003903; UTM.
DR Pfam: PF000036; efhand; 4.
DR SMART: SM00054; EFh; 3.
DR SMART: SM00027; EH; 3.
DR PROSITE: PS00018; EF_HAND; 2.
KW Phosphorylation; Calcium-binding; Repeat; Chromosomal translocation;
KW Proto-oncogene.
FT CA_BIND 173 184 EF_HAND 1 (POTENTIAL).
FT CA_BIND 236 247 EF_HAND 2 (POTENTIAL).
FT DOMAIN 599 827 15 X 3 AA REPEATS OF D-P-F.
FT REPEAT 599 601 1.
FT REPEAT 623 625 1.
FT REPEAT 629 631 3.
FT REPEAT 634 636 4.
FT REPEAT 640 642 5.
FT REPEAT 645 647 6.
FT REPEAT 651 653 7.
FT REPEAT 664 666 8.
FT REPEAT 672 674 9.
FT REPEAT 692 694 10.
FT REPEAT 709 711 11.
FT REPEAT 737 739 12.
FT REPEAT 798 800 13.
FT REPEAT 804 806 14.
FT REPEAT 825 827 15.
FT DOMAIN 768 850 PRO-RICH.
FT MOD_RES 132 132 PHOSPHORYLATION (BY TYR-KINASES)
FT CONFLICT 822 822 M -> I (IN REF. 2).
FT SEQUENCE 896 AA; 98673 MW; A1B9PB04A07FABEB CRC64;

```

Query Match 14.3%; Score 154; DB 1; Length 896;
 Best Local Similarity 22.6%; Pred. No. 0.08; Mismatches 60; Indels 80; Gaps 7;
 Matches 53; Conservative 42;

```

RESULT 22
ID MYH_HUMAN STANDARD; PRT; 1976 AA.
AC P35580;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, nonmuscle type B (cellular myosin heavy chain,
DE type B) (Nonmuscle myosin heavy chain-B) (NMHC-B).
GN MYH10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96025307; Pubmed=7499478;
RA Phillips C.L., Yamakawa K., Adelstein R.S.;
RT "Cloning of the cDNA encoding human nonmuscle myosin heavy chain-B and
RT analysis of human tissues with isoform-specific antibodies.";
RL J. Muscle Res. Cell Motil. 16:379-389(1995).
RN [2]
RP SEQUENCE OF 63-722 FROM N.A.
RX MEDLINE=91316803; Pubmed=1860190;
RA Simons M., Wang M., McBride O.W., Kawamoto S., Yamakawa K.,
RA Gula D., Adelstein R.S., Weir L.;
RT "Human nonmuscle myosin heavy chains are encoded by two genes located
RT on different chromosomes.";
RL Circ. Res. 69:530-539(1991).
CC -1- FUNCTION: CELLULAR MYOSIN APPEARS TO PLAY A ROLE IN CYTOKINESIS,
CC CELL SHAPE, AND SPECIALIZED FUNCTIONS SUCH AS SECRETION AND
CC CAPING.
CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY
CC CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2
CC REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M69181; AAA99177.1; -.
DR PIR: B61231; B61231.
DR HSSP: P08799; 1LVK.
DR MIM: 160776; -.
DR InterPro: IPR000048; IQ.
DR InterPro: IPR000409; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR002017; Spectrin.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ; 1.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MYSC; 1.
DR PROSITE: PS50096; IQ; 1.
KW Myosin; ATP-binding; Calmodulin-binding; Actin-binding;
KW Coiled coil; Alkylation; Multigene family.
FT DOMAIN 1 785 MYOSIN HEAD-LIKE.
FT DOMAIN 786 815 IQ.

```


FT DOMAIN 845 1976 COILED COIL (POTENTIAL).
 FT NE_BIND 178 185 ATP (POTENTIAL).
 FT MOD_RES 701 701 ALKYLIATION (SH-1) (POTENTIAL).
 FT MOD_RES 711 711 ALKYLIATION (SH-2) (POTENTIAL).
 SQ SEQUENCE 1976 AA, 228938 MW, B2BB87FF35EAL24F CRC64;

Query Match 14.3%; Score 154; DB 1; Length 1976;
 Best Local Similarity 21.0%; Pred. No. 0.18;
 Matches 64; Conservative 58; Mismatches 93; Indels 90; Gaps 10;

QY 4 INKLFPLAEEENV-----LDAEF--LKNEIDSVKAOISQK 38
 DB 1078 IDEKLTQIAKKEELQALARGDEDTLKKNNALKVRELQAOIAELQEDFESEKASRNKA 1137
 QY 39 DREKRSQAIIDTLDLTLEERNATVESLQNAKNAKMLCSLTK-----QMKFL 87
 DB 1138 EKQKRDLSSELEALKTLELDLTDTTAAQQLRTKREQVVALKKALEETKNNHQAQIDM 1197
 QY 88 EORQ----DETKQAREEAHRLK-----CKMKTMEQIELLIQSOR--- 122
 DB 1198 RQRATALEELSEQLDEQAKFKFKALEKNKQGLDNTKELACEVAVLQOVAESHKRKKL 1257
 QY 123 -SEVEEIRDMGVQSAVEQLAVYCVSLKKEYEN---LKEARRA----- 162
 DB 1258 DAQVOELHAYVEGDRRLRVELAEKASKLQNELDNVSTLEEAKEKGIKFAKDAASLESQ 1317
 QY 163 --TGEGLADRLKDDVSSRSKTKTLNTE---LDQAKLELRSAQKQLO---SADDEITSLR 213
 DB 1318 QDTQELQOETTRKLNLSRIROLEEKNLSIQEQEEEEKAKNMLEKQVIALQSOLADTK 1377
 QY 214 KKSDD 218
 DB 1378 KKVD 1382

RESULT 23
 ID MYSD_CAEEL STANDARD; PRT; 1938 AA.
 AC P02567; 019674;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin heavy chain D (MHC D).
 GN MYO-1 OR R06C7.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabdilitoidea;
 OC Rhabdilitidae; Peloderiinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=89178677; PubMed=2926820;
 RA Dbb N.D., Maruyama I.N., Krause M., Karn J.;
 RT "Sequence analysis of the complete Caenorhabditis elegans myosin
 heavy chain gene family.";
 RL J. Mol. Biol. 205:603-613(1989).
 RN [2]
 RP SEQUENCE OF 34-1795 FROM N.A.
 RX MEDLINE=83273600; PubMed=6576334;
 RA Karn J., Brenner S., Barnett L.;
 RT "Protein structural domains in the Caenorhabditis elegans unc-54
 myosin heavy chain gene are not separated by introns.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:4253-4257(1983).
 RN [3]
 RP SEQUENCE OF 115-365 AND 1492-1763 FROM N.A.
 RX MEDLINE=85201409; PubMed=3888374;
 RA Karn J., Dbb N.J., Miller D.M.;
 RT "Cloning nematode myosin genes.";
 RL Cell Muscle Motil. 6:185-237(1985).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;

RA Gardner A., McMurray A.;
 RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MUSCLE CONTRACTION.
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (LMC)
 CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MHC-2).
 CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -1- TISSUE SPECIFICITY: FOUND EXCLUSIVELY IN THE PHARYNGEAL MUSCLE.
 CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -1- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN
 CC C.ELEGANS.
 CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X08065; CAA30854.1; -;
 DR EMBL; M37232; AAA28119.1; -;
 DR EMBL; M37234; AAA28120.1; -;
 DR EMBL; Z71266; CAA95848.1; -;
 DR EMBL; Z71261; CAA95848.1; JOINED.
 DR EMBL; Z71261; CAA95806.1; -;
 DR EMBL; Z71266; CAA95806.1; JOINED.
 DR PIR; S02772; MMKW1.
 DR HSSP; P08799; 1MND.
 DR WormBep; R06C7.10; CE06253.
 DR InterPro; IPR004009; Myosin_N.
 DR InterPro; IPR002928; Myosin_tail.
 DR InterPro; IPR001609; myosin_head.
 DR Pfam; PF00063; myosin_head; 1.
 DR Pfam; PF02736; Myosin_N; 1.
 DR Pfam; PF01576; Myosin_tail; 1.
 DR PRINTS; PR00193; MYOSINHEAVY.
 DR PRODom; PD000355; myosin_head; 1.
 DR SMART; SM00242; MYSC; 1.
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KW ATP-binding; Methylation; Alkylation; Multigene family.
 FT FT 1 845
 FT DOMAIN 846 1938
 FT DOMAIN 846 1170
 FT DOMAIN 1171 1938
 FT DOMAIN 846 1938
 FT NP_BIND 177 184
 FT DOMAIN 660 682
 FT DOMAIN 764 778
 FT MOD_RES 128 128
 FT MOD_RES 700 700
 FT MOD_RES 710 710
 FT CONFLICT 94 94
 FT CONFLICT 98 98
 FT CONFLICT 377 377
 FT CONFLICT 389 390
 FT CONFLICT 391 391
 FT CONFLICT 408 408
 FT CONFLICT 474 474
 FT CONFLICT 577 577
 FT CONFLICT 681 681
 FT CONFLICT 1373 1373
 FT CONFLICT 1659 1659
 SQ SEQUENCE 1938 AA, 223255 MW, 387399C8F63A4CFA CRC64;
 V -> D (IN REF. 2).
 V -> D (IN REF. 4).
 W -> N (IN REF. 2).
 W -> G (IN REF. 2).
 L -> F (IN REF. 4).
 I -> N (IN REF. 4).
 S -> D (IN REF. 2).
 E -> Q (IN REF. 3).
 RODLIKE TAIL (S2 AND LM DOMAINS).
 ALPHA-HELICAL TAILPEPCE (SHORT S2).
 LIGHT MEROMYOSIN (LM).
 COILED COIL (POTENTIAL).
 ATP (BY SIMILARITY).
 ACTIN-BINDING.
 METHYLATION (TRI-).
 ALKYLIATION (SH-1).
 ALKYLIATION (SH-2).
 F -> E (IN REF. 2).
 A -> R (IN REF. 2).
 V -> D (IN REF. 4).
 DV -> GD (IN REF. 2).
 V -> D (IN REF. 4).
 W -> N (IN REF. 2).
 W -> G (IN REF. 2).
 L -> F (IN REF. 4).
 I -> N (IN REF. 4).
 S -> D (IN REF. 2).
 E -> Q (IN REF. 3).

```

Query Match      14.3%; Score 153.5; DB 1; Length 1938;
Best Local Similarity 23.5%; Pred. No.0.19; Indels 67; Gaps 11;
Matches 63; Conservative 56; Mismatches 82;

OY 15 EENVLDAEFLKNELDVYKQALSOQDRE-----KRDQAIIIDTLRDLTEERNA---TV 63
DQ 923 EKSINDANDRLSEHEKKNADLEKQRRKQOQEVENLKRSTAEVNDGLAKSLSEKKAKEQOI 982
OY 64 ESLONALKAMLCSTLKKOKKFLDQ-----RODETQAREHARLCKM-KTMEQ 113
DQ 983 HSLDDENMSODETIGTKINKERKLEENNRQALVDLQAEAKQA--QANRLRGKLEQITLDE 1040
OY 114 IELLQSO---RSEVEEMIR---DMVGQSAVQOLAVYCV---SLKKEENL----- 156
DQ 1041 MEAVEERKRIRATEKSKRRVSEGLKGAQETIDELSAIKETASLKKKADHIALGVR 1100
OY 157 -----KEARKATGELADRLKKDLVSSRSK-----DKTINTELDQ 190
DQ 1101 IEDBOALANRLTROSKENAQRIEIEDELEHER-QSRSKADRAVARELQRELDENRLDE 1159
OY 191 AKLELRSQKQLOASADOEITSLKSSD 218
DQ 1160 QNKQLEIQDQNNKKRSEITKFRDDE 1187

RESULT 24
TMFL_HUMAN STANDARD; PRT; 1093 AA.
AC P82094;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE TATA element modulatory factor (TMF).
GN TMFL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. AND FUNCTION.
RC TISSUE=Cervical carcinoma;
MDL=93028466; PubMed=1409643;
RA Garcia J.A., Ou S.-H.I., Wu F., Lusis A.J., Sparkes R.S., Gaynor R.B.;
RT "Cloning and chromosomal mapping of a human immunodeficiency virus 1
RT 'TATA' element modulatory factor."
RL Proc. Natl. Acad. Sci. U.S.A. 89:9372-9376(1992).
CC -!- FUNCTION: THIS PROTEIN BINDS THE HIV-1 TATA ELEMENT AND INHIBITS
CC TRANSCRIPTIONAL ACTIVATION BY THE TATA-BINDING PROTEIN (TBP).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L01042; AAD54608.1; -.
DR MIM: 601126; -.
KW Transcription regulation; DNA-binding; Repressor; Coiled coil.
FT DOMAIN 439 922 COILED COIL (POTENTIAL).
FT DOMAIN 984 1092 COILED COIL (POTENTIAL).
SQ SEQUENCE 1093 AA; 123170 MW; 261338B5F4677BE CAC64;

Query Match      14.2%; Score 153; DB 1; Length 1093;
Best Local Similarity 22.6%; Pred. No.0.11;
Matches 53; Conservative 49; Mismatches 99; Indels 34; Gaps 5;

OY 11 LAQEEENVLDAEFLKNELDVYKQALSOQDREKRDQAIIIDTLRDL---ERNATVESIQ 67

```

```

DQ 441 LSEKEDYCKTVEFLNEKLEKREAOULLSKREKALLEAFONLKDMEFRVKESSSISLK 500
OY 68 N-----ALKRKMCLC---STLKKOKK-----PLEQRODETQARE 99
DQ 501 DEFQRIAEAKKQVQLCKERDAKKEITKNIKELATRLNSSETADLLKEDEQIRGLME 560
OY 100 EAHRLCKMKKTMEQIELLOSORSEVEEMIRDMVGQSAVQOLAVYCVSLKKEENLKEA 159
DQ 561 EGEKLSQOQLNSNIITKRLAKDKDENMVAKL---KKYKLEELQHLKQVLDGKEEV 617
OY 160 RKATGELADRLKKDLVSSRSKTLTNTLQDAKLELSAQKQDQADQETSLRK 214
DQ 618 EKQRENNIKKLNSVVERQEKDLGRLOVMDLEKKNRSIQALADSAVKELTDLK 672

RESULT 25
MYHD_HUMAN STANDARD; PRT; 1938 AA.
AC Q9UKX3; Q95252;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, skeletal muscle, extraocular (MYH13).
GN MYH13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Extraocular muscle;
MDL=99318869; PubMed=1038558;
RA Weiss A., Schiaffino S., Lehmann L.A.;
RT "Comparative sequence analysis of the complete human sarcomeric myosin
RT heavy chain family: Implications for functional diversity."
RL J. Mol. Biol. 290:61-75(1999).
RN [2]
RP SEQUENCE OF 1917-1938 FROM N.A.
RC TISSUE=Extraocular muscle;
MDL=99026150; PubMed=9806854;
RA Winters L.M., Briggs M.M., Schachar F.;
RT "The human extraocular muscle myosin heavy chain gene (MYH13) maps to
RT the cluster of fast and developmental myosin genes on chromosome 17."
RL Genomics 54:188-189(1998).
CC -!- FUNCTION: MUSCLE CONTRACTION.
CC -!- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF111782; AAD29948.1; -.
DR EMBL: AF075248; AAC83241.1; -.
DR HSSP: P08799; 1MND.
DR MIM: 603487; -.

```

DR InterPro: IPR000048; IQ.
 DR InterPro: IPR000409; Myosin_N.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IQ: 2.
 DR Pfam: PF00063; myosin_head; 1.
 DR Pfam: PF02736; Myosin_N; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRODOM: PD000355; myosin_head; 1.
 DR SMART: SM00015; IQ: 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS50096; IQ; 1.
 DR Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KW Calmodulin-binding; ATP-binding; Methylation; Alkylation;
 KW Multigene family.
 KW MYOSIN HEAD-LIKE.
 FT DOMAIN 1 784
 FT DOMAIN 785 814
 FT DOMAIN 843 1938
 FT NP_BIND 179 186
 FT DOMAIN 659 681
 FT DOMAIN 761 775
 FT MOD_RES 130 130
 FT MOD_RES 699 699
 FT MOD_RES 709 709
 SQ SEQUENCE 1938 AA; 223678 MW; 1F6D006416381CD5 CRC64;

Query Match 14.28; Score 153; DB 1; Length 1938;
 Best Local Similarity 24.9%; Pred. No. 0.2;
 Matches 59; Conservative 40; Mismatches 74; Indels 64; Gaps 9;

OY 20 DAEFLKNEIDSVKAKLSQKREKRSQAIIID---TLRDLTEENNAVESIQNALNAEM 75
 DB 1382 DAIO RTELEBAKKKLRQLQEAKEKETETANSKCSLEKTKRQLQGEVEDLMRLERSHT 1441
 OY 76 LCSTL-KKQKF-----LEQRDETKQAREAHRLCKM-----KTMEQIEL 116
 DB 1442 ACATLDKKNRFDKVLAEWKQKDESSQALEAQSRLSTELFKRNVAEEVVDQLET 1501
 OY 117 L-----LQSORSEVEEMIRDMGYGQSAVEQLAVYCVSLKKEYENLKARRKATGELADRL 170
 DB 1502 LRRENKMIQOEISDLTEQIETGTG-----KNLQEAER-TKKLYEOE 1540
 OY 171 KKD-----LVSSRSKLTNTLMDQAKLELSAQKQDLSAQDEITSLEKKS 216
 DB 1541 KSDLQVALEEVGSLHEBESKILRVQLLEQVSKSEL--DKRTIEKDEIEQLKRN 1594
 RESULT 26
 MYSA_CAEEL
 ID MYSA_CAEEL STANDARD; PRT; 1969 AA.
 AC P12844;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, last sequence update)
 DT 16-OCT-2001 (Rel. 40, last annotation update)
 DE Myosin heavy chain A (MHC A).
 GN MYO-3
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RX MEDLINE=89178677; PubMed=2926820;
 RA Dibb N.J., Maruyama I.N., Krause M., Karn J.;
 RT "Sequence analysis of the complete Caenorhabditis elegans myosin
 heavy chain gene family.";
 RL J. Mol. Biol. 205:603-613(1989).
 CC -!- FUNCTION: MUSCLE CONTRACTION.
 CC -!- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)

CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
 CC -!- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
 CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
 CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
 CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -!- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
 CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
 CC -!- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
 CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 CC SUBFRAGMENT (S2).
 CC -!- MISCELLANEOUS: THERE ARE FOUR DIFFERENT MYOSIN HEAVY CHAINS IN
 CC C.ELEGANS.
 CC -!- MISCELLANEOUS: MHC A AND MHC B ARE FOUND EXCLUSIVELY IN THE BODY
 CC WALL MUSCLE. THEY CO-ASSEMBLE INTO BODY WALL THICK FILAMENT.
 CC -!- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 IQ DOMAIN.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X08067; CA30856.1; -.
 DR PTR: S02771; S02771.
 DR HSP: P08799; ILVK.
 DR InterPro: IPR000048; IQ.
 DR InterPro: IPR000409; Myosin_N.
 DR InterPro: IPR002928; Myosin_tail.
 DR InterPro: IPR001609; myosin_head.
 DR Pfam: PF00612; IQ: 1.
 DR Pfam: PF00063; myosin_head; 1.
 DR Pfam: PF02736; Myosin_N; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRODOM: PD000355; myosin_head; 1.
 DR SMART: SM00015; IQ: 1.
 DR SMART: SM00242; MYSC; 1.
 DR PROSITE: PS50096; IQ; 1.
 DR Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KW ATP-binding; Methylation; Alkylation; Multigene family.
 KW MYOSIN HEAD-LIKE.
 FT DOMAIN 1 793
 FT DOMAIN 794 823
 FT DOMAIN 857 1969
 FT NP_BIND 179 186
 FT DOMAIN 667 689
 FT DOMAIN 770 784
 FT MOD_RES 130 130
 FT MOD_RES 707 707
 FT MOD_RES 717 717
 SQ SEQUENCE 1969 AA; 225509 MW; 64577BBAF7EAD80A CRC64;

Query Match 14.18; Score 152; DB 1; Length 1969;
 Best Local Similarity 24.38; Pred. No. 0.23;
 Matches 62; Conservative 47; Mismatches 98; Indels 48; Gaps 8;

OY 5 NKLFEFLQAEENVLDAEFLKNEIDSVKAKLSQK-----DREKRS 45
 DB 898 NALFLSLETEKANLADAEERNEKLNQKATLESLSLDTQGLEDMQERNEDLAQKKKTD 957
 OY 46 QAITDT-----LRDTLEF---RNATVESIQNALNAEMLCSTLKKOMFLERQD 92
 DB 958 QELSDTQKHVQDLELSRKAEQEKOSRDHNIRSDDEMANODEAVARLNKREK---HQEE 1014
 OY 93 ETKQAREAHRLCKMKTMEQIELLQSORSEVEEMI-----RDMGYGQSAVEQLAVYCVS 148
 DB 1015 SNRKLNEIDQSEEDKVNHLKTRNKLEQOMDELENDIREKRSKGTIEKKRKYE---GD 1071
 OY 149 LKKEYENLKARRKATGELADRLKK---DLVSSRSKLTNTLT---ELDQAKLELSAQKDL 202

Db 1072 LKVAENIDEITKQKHVETLTKREEDLHTHTNAKLAEINSILAKQLRLIELFARNNEL 1131
 QY 203 QSAOETSLEKSD 217
 Db 1132 EEELEARNRSKSD 1146

RESULT 27
 H1P1_HUMAN
 ID H1P1_HUMAN STANDARD; PRT; 995 AA.
 AC 000291; 000328;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Huntingtin interacting protein 1 (H1P-1) (Fragment).
 GN H1P1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hug A.H.M., Nichol K., Osborne L., Scherer S.W., Squitieri F.,
 RA Hayden M.R.;
 RT "Genomic organization of the human H1P1 gene and its exclusion as a
 RT candidate gene in a family diagnosed with Huntington disease without
 RT CAG expansion." to the EMBL/GenBank/DBJ databases.
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 82-995 FROM N.A.
 RX MEDLINE=97285121; PubMed=9140394;
 RA Kalchman M.A., Kolde H.B., McCutcheon K., Graham R.K., Nichol K.,
 RA Nishiyama K., Kazeml-Esfarjani P., Lynn F.C., Wellington C.,
 RA Metzler M., Goldberg Y.P., Kanazawa I., Geitz R.D., Hayden M.R.;
 RT "H1P1, a human homologue of S. cerevisiae Sla2p, interacts with
 RT membrane-associated huntingtin in the brain."
 RL Nat. Genet. 16:44-53(1997).
 RN [3]
 RP SEQUENCE OF 82-159 FROM N.A.
 RA Bradshaw H., Hinds K., Harrison M.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 203-602 FROM N.A.
 RC TISSUE-Brain;
 RX MEDLINE=97227296; PubMed=9147654;
 RA Wanker E.E., Royra C., Scherzinger E., Hasenbank R., Waelter S.,
 RA Taft D., Colicelli J., Lehnach H.;
 RT "H1P-1: a huntingtin interacting protein isolated by the yeast two-
 RT hybrid system."
 RL Hum. Mol. Genet. 6:487-495(1997).
 RN [5]
 RP ALTERNATIVE SPLICING.
 RX MEDLINE=20515263; PubMed=11063258;
 RA Chopra V.S., Metzler M., Rasper D.M., Engqvist-Goldstein A.E.Y.,
 RA Singaraja R., Gan L., Fichter K.M., McCutcheon K., Drubin D.,
 RA Nicholson D.W., Hayden M.R.;
 RT "H1P2 is a non-proapoptotic member of a gene family including H1P1,
 RT an interacting protein with huntingtin."
 RL Mamm. Genome 11:1006-1015(2000).
 CC -1- FUNCTION: MAY PLAY A FUNCTIONAL ROLE IN THE CELL FILAMENT
 CC NETWORKS.
 CC -1- SUBUNIT: BINDS HUNTINGTIN. THIS INTERACTION IS RESTRICTED TO THE
 CC BRAIN.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. MEMBRANE-ASSOCIATED PROTEIN.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; H1P1-1 AND H1P1-2. ARE PRODUCED
 CC BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: UBICITOUSLY EXPRESSED WITH THE HIGHEST LEVEL
 CC IN BRAIN.
 CC -1- MISCELLANEOUS: THE AFFINITY OF THE HUNTINGTIN PROTEIN-H1P1
 CC INTERACTION IS INVERSELY CORRELATED TO THE LENGTH OF THE
 CC POLYGLUTAMINE TRACT ADDED TO THE HUNTINGTIN PROTEIN IN HUNTINGTON
 CC DISEASE.

CC -1- SIMILARITY: BELONGS TO THE SLA2 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC EMBL; AF052288; AAC33564.1; -
 DR EMBL; AF052288; AAC33564.1; JOINED.
 DR EMBL; AF052261; AAC33564.1; JOINED.
 DR EMBL; AF052262; AAC33564.1; JOINED.
 DR EMBL; AF052263; AAC33564.1; JOINED.
 DR EMBL; AF052264; AAC33564.1; JOINED.
 DR EMBL; AF052265; AAC33564.1; JOINED.
 DR EMBL; AF052266; AAC33564.1; JOINED.
 DR EMBL; AF052267; AAC33564.1; JOINED.
 DR EMBL; AF052268; AAC33564.1; JOINED.
 DR EMBL; AF052269; AAC33564.1; JOINED.
 DR EMBL; AF052270; AAC33564.1; JOINED.
 DR EMBL; AF052271; AAC33564.1; JOINED.
 DR EMBL; AF052272; AAC33564.1; JOINED.
 DR EMBL; AF052273; AAC33564.1; JOINED.
 DR EMBL; AF052274; AAC33564.1; JOINED.
 DR EMBL; AF052275; AAC33564.1; JOINED.
 DR EMBL; AF052276; AAC33564.1; JOINED.
 DR EMBL; AF052277; AAC33564.1; JOINED.
 DR EMBL; AF052278; AAC33564.1; JOINED.
 DR EMBL; AF052279; AAC33564.1; JOINED.
 DR EMBL; AF052280; AAC33564.1; JOINED.
 DR EMBL; AF052281; AAC33564.1; JOINED.
 DR EMBL; AF052282; AAC33564.1; JOINED.
 DR EMBL; AF052283; AAC33564.1; JOINED.
 DR EMBL; AF052284; AAC33564.1; JOINED.
 DR EMBL; AF052285; AAC33564.1; JOINED.
 DR EMBL; AF052286; AAC33564.1; JOINED.
 DR EMBL; AF052287; AAC33564.1; JOINED.
 DR EMBL; U79734; AAC51257.1; -
 DR EMBL; AAC004491; AAC08319.1; -
 DR EMBL; Y09420; CAA70574.1; -
 DR MIM; 601767; -
 DR InterPro: IPR001026; ENTH.
 DR InterPro: IPR002558; ILMEQ.
 DR Pfam: PF01407; ENTH. 1.
 DR Pfam: PF01608; ILMEQ. 1.
 DR ProDom: PD011820; ILMEQ. 1.
 DR SMART: SM00273; ENTH. 1.
 DR SMART: SM00307; ILMEQ. 1.
 KW Actin-binding; Coiled coil.
 FT NON_TER 1 1
 FT DOMAIN 326 602 COILED COIL (POTENTIAL).
 FT DOMAIN ? ? TALIN-LIKE.
 FT CONFLICT 203 209 KLHSCP -> EPPAAS (IN REF. 4).
 FT CONFLICT 597 602 LNLEE -> STBPRI (IN REF. 4).
 SQ SEQUENCE 995 AA; 111633 MW; 72BDAA16AFD1C40 CRC64;

Query Match 14.1%; Score 151.5; DB 1; Length 995;
 Best Local Similarity 25.1%; Pred. No. 0.12;
 Matches 61; Conservative 56; Mismatches 95; Indels 31; Gaps 10;

QY 4 INKFFPLAOE---ENNVADEFKLNELDSVKAQLSKDKREKDSQAIIITLRLTEER 59
 Db 367 VSELEADLAQOHLRQADADCEFLRAELDELRLQRDRTEKAQR-SLSETERKAQANEOR 425

QY 60 NATVESLONAL--NKAEML--CSFLKQMKFLBORO---DEFGQAREAHRL--KCKMK 109
 Db 426 YSKLKEKYSVLQNHADLKLKNAEVTQVSMARQAVDLEREKKELEDSLEKSDQSKRK 485

QY 110 TMEQIEILLQSORSEVEEMIRDMGVGSAVE-----QLAVYCVSLKKEYENLKARKA 162
 Db 486 TQEQLE-VLESLSKQELATSGRELQVLQGLSLETSNQSEANMAAFEALEKREDSLVSCAH 544

```

Qy 163 TGLADRLKKDLVSSRSKLTINTEL-----DQAKLELSRSKQKLDQADQETSLRRKSD 217
Db 545 REELSALRKRELODTQLKLASTEESMCKOLAKDKMKLLVSRK---AAEYVYDALNMLE 601
Qy 218 DPP 220
Db 602 EPP 604

RESULT 28
ID RA50_PYPAB STANDARD; PRT; 880 AA.
AC Q9U2C8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE DNA double-strand break repair rad50 ATPase.
GN RAD50 OR PAB0812.
OS Pyrococcus abyssi.
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=29292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ORSAY;
RA Hellig R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
structure and evolution."
CC -1- FUNCTION: Involved in DNA double-strand break repair (DSBR). The
rad50/mre11 complex possesses single-strand endonuclease activity
and ATP-dependent double-strand-specific exonuclease activity.
rad50 provides an ATP-dependent control of mre11 by unwinding
and/or repositioning DNA ends into the mre11 active site (By
similarity).
CC -1- SUBUNIT: Forms a complex with mre11 (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL: AJ248286; CAB50131.1; -.
DR InterPro: IPR003593; AAA.
DR InterPro: IPR003439; ABC_transportr.
DR InterPro: IPR001687; ATP_GTP_A.
DR InterPro: IPR001238; RecF.
DR Pfam: PF00470; RecF. 1.
DR SMART: SM00382; AAA; 1.
KW DNA repair; Hydrolase; ATP-binding; Coiled coil; Complete proteome.
FT NP_BIND 30 37 ATP (BY SIMILARITY).
FT DOMAIN 144 745 COILED COIL (POTENTIAL).
SQ SEQUENCE 880 AA; 103970 MW; FDB177EC7E026479 CRC64;

Query Match 14.0%; Score 151; DB 1; Length 880;
Best Local Similarity 19.4%; Pred. No. 0.11;
Matches 54; Conservative 57; Mismatches 72; Indels 96; Gaps 8;

Qy 10 DLADBEENVLDLAEFLKLNLDLVKQSLQKDKREKDSQAIIITLTKDLERANATVESL--- 66
Db 226 ELEVENRENVKLEISIKISIKISLQVEKLGKRGKLEKIYQIRISIEKRAKISELEEI 285
Qy 67 -----ONALNKRAEMLCSTLKKOMKPLE----- 88
Db 286 VKDIPKLOEKREYKRLKGFDEYESKRLKLEKSKSESLKALEYIKGEKKERAE 345
Qy 89 -----ORODETKQAREAHRLCKMKTM--EQIEEL--LQSGRS 123

```

```

Db 346 EIREKLEISIEKRLKLELPPVEELEDAAKQVOKTERLKLKGLSPGEVIEKLESEKERT 405
Qy 124 EVEEMIRDM--GVGQSAVEDOLAVYCVSLKKEVLEKRRKATG-----ELDRKDKDL 174
Db 406 EIEPAKEITTRIGOMOEK-----NEKMAIEELRRAKGCPVCGRELTIEHKEL 457
Qy 175 VSSRSKLTINTELQAKLELSRSKQKLDQADQETSLR 213
Db 458 -----MERTLEIKKEIEELKRTTEERKLR 483

RESULT 29
ID MYSP_CAEEL STANDARD; PRT; 882 AA.
AC P10567;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Paramyosin.
GN UNC-15.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=BRISTOL N2;
RX MEDLINE=89329026; PubMed=2754728;
RA Kagawa H., Gengyo K., McLachlan A.D., Brenner S., Karn J.;
RT "Paramyosin gene (unc-15) of Caenorhabditis elegans. Molecular
cloning, nucleotide sequence and models for thick filament
structure."
RT J. Mol. Biol. 207:311-333(1989).
RN [2]
RP ERRATUM.
RA Kagawa H., Gengyo K., McLachlan A.D., Brenner S., Karn J.;
RL J. Mol. Biol. 211:665-665(1990).
RN [3]
RP PHOSPHORYLATION.
RX MEDLINE=89329036; PubMed=2754733;
RA Schriefer L.A., Waterson R.H.;
RT "Phosphorylation of the N-terminal region of Caenorhabditis elegans
RT paramyosin."
RL J. Mol. Biol. 207:451-454(1989).
CC -1- FUNCTION: PARAMYOSIN IS A MAJOR STRUCTURAL COMPONENT OF
MANY THICK FILAMENTS ISOLATED FROM INVERTEBRATE MUSCLES.
CC
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC
CC -1- DOMAIN: FOR MOST OF ITS LENGTH, PARAMYOSIN APPEARS TO FORM AN
ALPHA-HELICAL COILED COIL AND SHOWS THE HEPTAD REPEAT OF
HYDROPHOBIC AMINO ACID RESIDUES AND THE 28-RESIDUE REPEAT OF
CHARGED AMINO ACIDS CHARACTERISTIC OF MYOSIN HEAVY CHAINS.
CC HOWEVER, PARAMYOSIN DIFFERS FROM MYOSIN IN HAVING NON-HELICAL
EXTENSIONS AT BOTH TERMINI AND AN ADDITIONAL "SKIP" RESIDUE THAT
INTERRUPTS THE 28-RESIDUE REPEAT. THE DISTRIBUTION OF CHARGED
RESIDUES IS ALSO DIFFERENT FROM MYOSIN HEAVY CHAINS.
CC -1- PTM: PHOSPHORYLATED ON SERINE RESIDUES IN THE NON-ALPHA-HELICAL
N-TERMINAL REGION.
CC
CC -1- SIMILARITY: HIGH TO MYOSIN HEAVY CHAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL: X08068; CA430857.1; -.
DR PIR: S04027; S04027.
DR HSSP: P80220; IDIP.
DR InterPro: IPR002928; Myosin_tail.
DR Pfam: PF01576; Myosin_tail; 1.

```

KW Coiled coil; Muscle protein; Thick filament; Myosin; Phosphorylation.
 FT DOMAIN 1 47 NONHELICAL REGION (POTENTIAL).
 FT DOMAIN 48 871 COILED COIL (POTENTIAL).
 FT DOMAIN 872 882 NONHELICAL REGION (POTENTIAL).
 FT DISULFID 143 143 INTERCHAIN (POTENTIAL).
 FT DISULFID 632 632 INTERCHAIN (POTENTIAL).
 SO SEQUENCE 882 AA; 101949 MW; 894C4BFF4A670C5F CRC64;

Query Match 14.0% Score 150.5; DB 1; Length 882;
 Best Local Similarity 24.4% Pred. No. 0.12; Mismatches 83; Indels 63; Gaps 10;
 Matches 63; Conservative 49; Mismatches 83; Indels 63; Gaps 10;

QY 6 KLFFDLQAEENVDAEFLKNEKLSVKAQLS-----QKDKREKRS----- 45
 DB 447 KLHDLHAKKALADANKLHDLNRLAGEIRELOTALKENDACQORDNENRQRLA 506
 QY 46 --QAIIDTLRDTLEBRNATVESLQNALN-----KADMLCSTLKK--- 82
 DB 507 ELQALRIMERRLQKEEEMELKRLNLFQFIDRLIALADAEARKSEI--SRLLKKYQA 564
 QY 83 QMKFLQKQDQETKQAREAHRLKCKMKTMEQIELL--LQSQRSVEEMIMDMVGQSAV 139
 DB 565 EIAELEMTVDLNRANITFAQ--KTIKKQEQKLTQLQASLEDTQRLQGVLDQYALAKRV 622
 QY 140 EQLAVYCVSLKREYENLEAKR-----ATGELAD--RLKKDIVSSRSKLTLELD 189
 DB 623 AALSALEECTALDNARAKRQAEVDELEANGRISDISINNLTSTKNTLE---TFLS 679
 QY 190 QAKLELSAQKDLQSDQ 207
 DB 680 TAQADLDEVTKELHNADE 697

RESULT 30
 SCPI_HUMAN STANDARD; PRT; 976 AA.
 ID SCPI_HUMAN
 AC 015431; 014963;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Synaptonemal complex protein 1 (SCP-1 protein).
 GN SCPI OR SCPI.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=97224467; PubMed=9119375;
 RA Meuwissen R.L.J., Meerts I., Hoovers J.M.N., Leschot N.J.,
 RA Heyting C.,
 RT "Human synaptonemal complex protein 1 (SCPI): isolation and
 RT characterization of the cDNA and chromosomal localization of the
 RT gene";
 RL Genomics 39:377-384(1997).
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=98037449; PubMed=9371398;
 RA Kondoh N., Nishina Y., Tsuchida J., Koga M., Tanaka H., Uchida K.,
 RA Inazawa J., Taketo M., Nozaki M., Nojima H., Matsuiya K., Namiki M.,
 RA Okuyama A., Nishimura Y.;
 RT "Assignment of synaptonemal complex protein 1 (SCPI) to human
 RT chromosome 1p13 by fluorescence in situ hybridization and its
 RT expression in the testis";
 RL Cytogenet. Cell Genet. 78:103-104(1997).
 CC -1- FUNCTION: MAJOR COMPONENT OF THE TRANSVERSE FILAMENTS OF
 CC SYNAPTONEMAL COMPLEXES (SCS). FORMED BETWEEN HOMOLOGOUS
 CC CHROMOSOMES DURING MEIOTIC PROPHASE.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR. IN TRIPARTITE SEGMENTS OF
 CC SYNAPTONEMAL COMPLEXES, BETWEEN LATERAL ELEMENTS IN THE NUCLEUS.

CC FOUND ONLY WHERE THE CHROMOSOME CORES ARE SYNAPSED. ITS N-TERMINUS
 CC IS FOUND TOWARDS THE CENTRE OF THE SYNAPTONEMAL COMPLEX WHILE THE
 CC C-TERMINUS EXTENDS WELL INTO THE LATERAL DOMAIN OF THE
 CC SYNAPTONEMAL COMPLEX (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: TESTIS.
 CC -1- DOMAIN: CONSISTS OF AN ALPHA-HELICAL STRETCH OF 700 AA RESIDUES,
 CC FLANKED BY N- AND C-TERMINAL GLOBULAR DOMAINS. THE C-TERMINAL
 CC DOMAIN HAS DNA-BINDING CAPACITY (BY SIMILARITY).

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; X95654; CA64956.1; -;
 DR EMBL; D67035; BAA22586.1; -;
 DR MIM; 602162; -;
 KW Nuclear protein; Meiosis; Cell division; Phosphorylation;
 KB DNA-binding; Coiled coil.

FT DOMAIN 12 100 ASF/GLU-RICH (ACIDIC).
 FT DOMAIN 107 798 COILED COIL (POTENTIAL).
 FT DOMAIN 117 120 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 679 682 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 880 883 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 961 969 ARG/LYS-RICH (BASIC).
 FT CONFLICT 46 106 L -> F (IN REF. 2).
 FT CONFLICT 106 106 F -> Y (IN REF. 2).
 FT CONFLICT 153 153 K -> C (IN REF. 2).
 FT CONFLICT 161 161 F -> T (IN REF. 2).
 FT CONFLICT 168 168 E -> D (IN REF. 2).
 FT CONFLICT 216 216 N -> S (IN REF. 2).
 FT CONFLICT 225 226 HG -> FE (IN REF. 2).
 FT CONFLICT 350 350 K -> N (IN REF. 2).
 FT CONFLICT 360 360 E -> D (IN REF. 2).
 FT CONFLICT 401 401 KN -> NY (IN REF. 2).
 FT CONFLICT 406 406 K -> I (IN REF. 2).
 FT CONFLICT 415 415 K -> T (IN REF. 2).
 FT CONFLICT 449 449 E -> D (IN REF. 2).
 FT CONFLICT 483 510 IQLATITTSQYYSKVEKDKLTLENEK -> YSYGHYHKW
 FT CONFLICT 516 528 TVLPKRGQRPKLSKRE (IN REF. 2).
 FT CONFLICT 516 528 LTFHCNKLSTLENK -> YFTLQOASPPPN (IN REF. 2).
 FT CONFLICT 549 549 N -> I (IN REF. 2).
 FT CONFLICT 560 560 K -> T (IN REF. 2).
 FT CONFLICT 805 805 E -> D (IN REF. 2).
 FT CONFLICT 941 941 P -> S (IN REF. 2).
 SO SEQUENCE 976 AA; 114069 MW; 8BA81D042AC2696B CRC64;

Query Match 14.0% Score 150; DB 1; Length 976;
 Best Local Similarity 26.2% Pred. No. 0.14; Mismatches 86; Indels 28; Gaps 9;
 Matches 59; Conservative 52; Mismatches 86; Indels 28; Gaps 9;

QY 10 DLQAEENVDAEFLKNEKLSVKAQLS-----QKDKREKRS----- 123
 DB 501 DLKTELEN-----EKLKTELTSHCNKLSLENKELTQETSDMTLEKQOEDINNNKQOE 556
 QY 68 NALNKAEMLCSTLKKQKFLQKQDQETKQAREAHRLKCKMKTMEQ---TELLSQRS 123
 DB 557 RMLQOINLQDETETQLNNELEYVREELKQRDE---VKCKLDKSEKCNMLRKQVERKNK 613
 QY 124 EYEM-----IRMGVGSQSAVEDOLAVYCVSLKREYENLEAKRATGELADRLKDIVSS 177
 DB 614 YIELQOENKALKKKGTAES--KOLNVEIKVNTLELELSAKCKKFEIDTYQKTEDEK 671
 QY 178 RSKTKLTNTELDQAKL---ELRSQKDLQSDQ-----EITSRLKK 215
 DB 672 KISEENLLEVEKAKVADAEVAVKLQKELDKRCQKHAEMVALMEK 716

```

RESULT 31
MYH8_HUMAN STANDARD: PRT; 1937 AA.
ID MYH8_HUMAN
AC P13535; O14910;
DT 01-JAN-1990 (Rel. 13, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, skeletal muscle, perinatal (MyHC-perinatal).
GN MYH8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=90323631; PubMed=2373371;
RA Karsch-Mizrachi I., Feghali R., Shows T.B. Jr., Leinwand L.A.;
RT "Generation of a full-length human perinatal myosin heavy-chain-
RL encoding cDNA."
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=95324556; PubMed=7601129;
RA Julian E.H., Kelly A.M., Pompidou A.J., Hoffman R., Schiaffino S.,
RA Stedman H.H., Rubinstein N.A.;
RT "Characterization of a human perinatal myosin heavy-chain
RL transcript."
RN [3]
RP SEQUENCE OF 502-1937 FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=90235862; PubMed=1691980;
RA Rober E., Buchberger-Seidl A., Braun T., Singh S., Goedde H.W.,
RA Arnold H.H.;
RT "Identification of three developmentally controlled isoforms of human
RL myosin heavy chains."
RN [4]
RP SEQUENCE OF 860-1937 FROM N.A.
RX MEDLINE=89234168; PubMed=2715179;
RA Feghali R., Leinwand L.A.;
RT "Molecular genetic characterization of a developmentally regulated
RL human perinatal myosin heavy chain."
RN [5]
RP SEQUENCE OF 1-46 FROM N.A.
RA Esser K., Tidhar A., Myszkowski M.;
RT "Isolation and characterization of the human perinatal MHC promoter."
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPEITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LM) AND 1 HEAVY MEROMYOSIN (HM). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entries requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M36769; AAC17185.1; -
DR EMBL; Z38193; CAA86293.1; -
DR EMBL; X51592; CAA35941.1; -
DR EMBL; M35250; AAA36346.1; -
DR EMBL; AF067143; AAC21557.1; -
DR PIR; A30220; A30220.
DR HSSP; P08799; ILVK.
DR MIM; 160741; -.
DR InterPro; IPR000048; IQ.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
KW Myosin; Muscle protein; coiled coil; thick filament; Actin-binding;
KW ATP-binding; Methylation; Alkylation; Multigene family;
KW Calmodulin-binding.
FT DOMAIN 1 780
FT MYOSIN HEAD-LIKE.
FT DOMAIN 781 813
FT IQ.
FT DOMAIN 842 1937
FT COILED COIL (POTENTIAL).
FT NP_BIND 181 188
FT ATP.
FT DOMAIN 658 680
FT ACTIN-BINDING.
FT DOMAIN 760 774
FT ACTIN-BINDING.
FT MOD_RES 132 132
FT METHYLATION (TR1-) (POTENTIAL).
FT MOD_RES 698 698
FT ALKYLATION (SH-1) (POTENTIAL).
FT MOD_RES 708 708
FT ALKYLATION (SH-2) (POTENTIAL).
FT CONFLICT 15 15
FT A -> R (IN REF. 2).
FT CONFLICT 970 970
FT E -> N (IN REF. 1 AND 4).
FT CONFLICT 1072 1072
FT M -> N (IN REF. 3).
FT CONFLICT 1247 1247
FT N -> H (IN REF. 1 AND 4).
FT CONFLICT 1251 1251
FT MC -> DG (IN REF. 3).
FT CONFLICT 1261 1261
FT E -> G (IN REF. 1 AND 4).
FT CONFLICT 1297 1297
FT K -> Q (IN REF. 1 AND 4).
FT CONFLICT 1377 1378
FT KY -> NT (IN REF. 3).
FT CONFLICT 1504 1505
FT EN -> AH (IN REF. 1 AND 4).
FT CONFLICT 1847 1847
FT E -> D (IN REF. 1 AND 4).
FT CONFLICT 1914 1914
FT D -> H (IN REF. 2).
SQ SEQUENCE 1937 AA; 222762 MW; A3EE2D151792E9E8 CRC64;

Query Match 14.0%; Score 150; DB 1; Length 1937;
Best Local Similarity 24.3%; Pred. No. 0.29;
Matches 64; Conservative 48; Mismatches 85; Indels 66; Gaps 9;

3 IIRKLFEDLAEEENVDLAEFLKNELDVSKAOLSKDREKRSQAIIDTLRDTLEERNAT 62
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1272 LINDL---TRQRRRLQTEAGEYSRQIDKDALVSQLSRSQASTQOIEELKHQLEETRA 1328
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
63 VESIQNALNKAEMLCSGLKQKMFLE-----ORDE 93
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1329 KNALAHNLQSSRRDCDLRLRQYEEQGAELQRLSKANSEVAQWRKYETDAIQRTTE 1388
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
94 TKRAREE-AHRLK-----CKKRTMQITILLIQSORSEVEEMIRPMGVQSAVEOL 142
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1389 LEEAKKKLQRLQLEAEHVEAVNAKCAASLEKTKQRLQ---NEVEDLMD-----VERS 1438
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
143 AVYCVSLKREYENILKEARKATGELADRLKLDIVSSRSKLTLTLELDQAK-----L 193
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1439 NACCAALDKQRNFDFVLSWMKKYETQALEASQKESLSSTELFKYKNVYESLDL 1498
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
94 E-LRSQKDLQSDQETTSLRK 215
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1499 ETLRRENKNLQ---QETSDLTEQ 1518
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```


CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D89992; BAA22069.1; -
 DR EMBL: D50476; BAA09069.1; -
 DR EMBL: DA3700; BAA07802.1; -
 DR HSSP: P08799; 1AMD,
 DR InterPro: IPR000048; IQ,
 DR InterPro: IPR004009; Myosin_N,
 DR InterPro: IPR002928; Myosin_tail,
 DR InterPro: IPR001609; myosin_head,
 DR Pfam: PF00612; IQ; 1,
 DR Pfam: PF00063; myosin_head; 1,
 DR Pfam: PF02736; Myosin_N; 1,
 DR Pfam: PF01576; Myosin_tail; 1,
 DR PRINTS: PR00193; MYOSINHEAVY,
 DR PRODOM: PD000355; myosin_head; 1,
 DR SMART: SM00015; IQ; 1,
 DR SMART: SM00242; MYSC; 1,
 DR PROSITE: PS00936; IQ; 1,
 KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
 KW ATP-binding; Methylation; Alkylation; Calmodulin-binding;
 KW Multigene family.
 FT DOMAIN 1 781 MYOSIN HEAD-LIKE.
 FT DOMAIN 782 811 IQ.
 FT DOMAIN 812 839 HINGE.
 FT DOMAIN 840 1935 COILED COIL (POTENTIAL).
 FT NP_BIND 178 185 ATP (POTENTIAL).
 FT DOMAIN 659 681 ACTIN-BINDING.
 FT DOMAIN 761 775 ACTIN-BINDING.
 FT MOD_RES 129 129 METHYLATION (SH-1).
 FT MOD_RES 699 699 ALKYLATION (SH-2).
 FT MOD_RES 709 709 ALKYLATION (SH-2).
 SQ SEQUENCE 1935 AA; 221599 MW; 9A1244B67D63C83B CRC64;

Query Match 13.98; Score 149; DB 1; Length 1935;
 Best Local Similarity 24.38; Pred. No. 0.33;

Matches 66; Conservative 52; Mismatches 84; Indels 70; Gaps 13;

QY 11 LAQEEENVLAELFLKNELDS-VKNAQLSQKDR-----KRDSQAITDLRLDLE-----DRNA 61
 DB 1552 LEHRESITILVQLELNQVKEIDIKLAEKDEBQIKRNSQRYVDSQSTLDSVRSRND 1611
 QY 62 TV---ESLQNALNKAEMLCSTLKKQ---MKFLEORODETRQAR---EAAHR---LKCK 107
 DB 1612 ALRVKKKMEGDLNEMETIQLSHANRQAALAKQLRNVOGLKDAQLHIDEAVRQGEDMKEQ 1671
 QY 108 MKTMEQJELLQSORSEVEEMIRMGVQSAVEQLAV-----YCSLKKY 153
 DB 1672 VAAVERNSLMDQAEIELRALAEQTERGRVAAVEQLDASERVGLHNSONTSLINTKKL 1721
 QY 154 E-----NLKFAKRA-----TGLADRLKKDYSS-----RSK-----L 181
 DB 1732 EADLVQVQGEVDADVQARNAAEKAKAITDAAMAAEELKKEQDTSAHLEKMKNLLEVTV 1791
 QY 182 KTLNTELDQAK-LELSAQKQLQASQDOETISL 212
 DB 1792 KDLQHRLEDAEASLAKMGKKQLQKLESVRREL 1823

RESULT 34
 MYH9_HUMAN
 ID MYH9_HUMAN STANDARD; PRT; 1960 AA.
 AC P35579; O60805;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Myosin heavy chain, nonmuscle type A (Cellular myosin heavy chain,
 type A) (Nonmuscle myosin heavy chain-A) (NMHC-A).

GN MYH9.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20057165; PubMed=10591208;
 RA Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,
 RA Ciamp M., Smink L.J., Alnsough R., Almeida J.P., Babbage A.,
 RA Baggaley C., Bailey J., Barlow K., Bates K.N., Beasley O., Bird C.P.,
 RA Blakey S., Bridgeman A.M., Buck D., Burgess J., Burrill W.D.,
 RA Burton J., Carder C., Carter N.P., Chen Y., Clark G., Clegg S.M.,
 RA Cobley V., Cole C.G., Collier R.E., Connor R.E., Conroy D., Corby N.,
 RA Coville G.J., Cox A.V., Davis J., Dawson E., Dhani P.D., Dockree C.,
 RA Dodswoth S.J., Durbin R.M., Ellington A., Evans K.L., Fey J.M.,
 RA Fleming K., French L., Garner A.A., Gilbert J.G.R., Goward M.E.,
 RA Graffham D., Griffiths M.N., Hall C., Hall R., Hall-Tamlyn G.,
 RA Heathcote R.W., Ho S., Holmes S., Hunt S.E., Jones M.C., Kershaw J.,
 RA Kimberley A., King A., Laird G.K., Langford C.F., Leversha M.A.,
 RA Lloyd C., Lloyd D.M., Martyn I.D., Mashreghi Mohammadi M.,
 RA Matthews L., Mccann O.T., Mccloy J., McLaren S., Mcmurray A.A.,
 RA Milne S.A., Mortimore B.J., Odell C.N., Pavitt R., Pearce A.V.,
 RA Pearson D., Phillimore B.J., Phillips S.H., Plumb R.W., Ramsay H.,
 RA Ramsey Y., Rogers L., Ross M.T., Scott C.E., Sehra H.K., Skuce C.D.,
 RA Smalley S., Smith M.L., Soderlund C., Spraggon L., Stewart C.A.,
 RA Sulston J.E., Swann R.M., Vaudin M., Wall M., Wallis J.M.,
 RA Whiteley M.N., Willey D., Williams L., Williams S., Williamson H.,
 RA Wilmer T.E., Wilming L., Wright C.L., Hubbard T., Bentley D.R.,
 RA Beck S., Rogers J., Shimizu N., Minoshima S., Kawasaki K., Sasaki T.,
 RA Asakawa S., Kudoh J., Shintani A., Shibuya K., Yoshizaki Y., Aoki N.,
 RA Mitsuyama S., Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S.,
 RA Do A., Do T., Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S.,
 RA Lai H., Lao H.T., Lewis J., Lewis S., Lin S.P., Lon P., Mala E.,
 RA Nguyen T., Pan H., Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S.,
 RA Sloan D., Song L., Wang Q., Wang Y., Wang Z., White J., Willingham D.,
 RA Wu H., Yao Z., Zhan M., Zhang G., Chissoe S., Murray J., Miller N.,
 RA Mux P., Fulton R., Johnson D., Bents G., Bentley D., Bradshaw H.,
 RA Bourne S., Cordes M., Du Z., Fulton L., Geola D., Graves T.,
 RA Hawkins J., Hinds K., Kemp K., Latreille P., Layman D., Ozerksy P.,
 RA Rohlfing T., Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K.,
 RA Nelson J., Korf I., Bedell J.A., Hillier L., Marais E., Waterston R.,
 RA Wilson R., Emanual B.S., Shaikh T., Kurahashi H., Salta S.,
 RA Budart M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
 RA Belmann L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,
 RA Peyrard M., Kedra D., Seroussi E., Franconi I., Tapia I., Bruder C.E.,
 RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
 RA Khan A.S., Lane L., Tiliahun Y., Wright H.;
 RL "The DNA sequence of human chromosome 22.";
 RL Nature 402:489-495(1999).
 RN [2]
 RP SEQUENCE OF 1-1337 FROM N.A.
 RX MEDLINE=92003925; PubMed=1912569;
 RA Toothaker L.E., Gonzalez D.A., Tung N., Lemons R.S., le Beau M.M.,
 RA Arnaout M.A., Clayton L.K., Tenen D.G.;
 RT "Cellular myosin heavy chain in human leukocytes: Isolation of 5'
 RT cDNA clones, characterization of the protein, chromosomal
 RT localization, and upregulation during myeloid differentiation.";
 RL Blood 78:1826-1833(1991).
 RN [3]
 RP SEQUENCE OF 1-715 FROM N.A.
 RX MEDLINE=91316803; PubMed=1860190;
 RA Simons M., Wang M., McBride O.W., Kawamoto S., Yamakawa K.,
 RA Gula D., Adelsstein R.S., Weir L.;
 RT "Human nonmuscle myosin heavy chains are encoded by two genes located
 RT on different chromosomes.";
 RL Circ. Res. 69:530-539(1991).
 RN [4]
 RP SEQUENCE OF 714-1960 FROM N.A.
 RX MEDLINE=90138958; PubMed=1967836;
 RA Saez C.G., Myers J.C., Shows T.B., Leinwand L.A.;
 RT "Human nonmuscle myosin heavy chain mRNA: generation of diversity
 RT through alternative polyadenylation.";

RL Proc. Natl. Acad. Sci. U.S.A. 87:1164-1168(1990).
RN [5]
RP VARIANT DFNA17 HIS-705.
RX MEDLINE-20489856; PubMed-11023810;
RA Lalwani A.K., Goldstein J.A., Kelley M.J., Luxford W., Castelein C.M.,
RA Mhatre A.N.,
RT "Human nonsyndromic hereditary deafness DFNA17 is due to a mutation in
RT nonmuscle myosin MYH9.";
RL Am. J. Hum. Genet. 67:1121-1128(2000).
RN [6]
RP VARIANTS MHA/ETNS/SBS K-93; C-702; C-1165; H-1424 AND K-1841.
RX MEDLINE-20428192; PubMed-10973259;
RA Seti M., Cusano M., Gangarossa S., Caridi G., Bordo D., Lo Nigro C.,
RA Gialleggeri G.M., Ravazzolo R., Savino M., Del Vecchio M., d'Apolito M.,
RA Iolascon A., Zelaneto L.L., Savoia A., Baldini C.L., Norris P.,
RA Magrini U., Belletti S., Heath K.E., Babcock M., Glucksmann M.J.,
RA Aliprandis E., Bizzaro N., Desnick R.J., Martignetti J.A.,
RT "Mutations in MYH9 result in the May-Hegglin anomaly, and Fechtner and
RT Sebastian syndromes.";
RL Nat. Genet. 26:103-105(2000).
RN [7]
RP VARIANTS MHA ILE-1155 AND LYS-1841.
RX MEDLINE-20428193; PubMed-10973260;
RA Kelley M.J., Jawien W., Ortel T.L., Korczak J.F.,
RT "Mutation of MYH9, encoding non-muscle myosin heavy chain A, in
RT May-Hegglin anomaly.";
RL Nat. Genet. 26:106-108(2000).
CC -1- FUNCTION: CELLULAR MYOSIN APPEARS TO PLAY A ROLE IN CYTOKINESIS,
CC CELL SHAPE, AND SPECIALIZED FUNCTIONS SUCH AS SECRETION AND
CC CAPRINC.
CC -1- SUBUNIT: MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2 HEAVY
CC CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC) AND 2
CC REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- DISEASE: DEFECTS IN MYH9 ARE THE CAUSE OF MAY-HEGGLIN ANOMALY
CC (MHA), AN AUTOSOMAL, DOMINANT MACROTHROMBOCYTOPENIA CHARACTERIZED
CC BY THROMBOCYTOPENIA, GIANT PLATELETS AND LEUCOCYTE INCLUSIONS.
CC -1- DISEASE: DEFECTS IN MYH9 ARE THE CAUSE OF FECHTNER SYNDROME
CC (ETNS), AN AUTOSOMAL, DOMINANT MACROTHROMBOCYTOPENIA CHARACTERIZED
CC WITH THROMBOCYTOPENIA, GIANT PLATELETS AND LEUCOCYTE INCLUSIONS.
CC WITH ADDITIONAL ALPORT-LIKE CLINICAL FEATURES OF SENSORINEURAL
CC DEAFNESS, CATARACTS AND NEPHRITIS.
CC -1- DISEASE: DEFECTS IN MYH9 ARE THE CAUSE OF SEBASTIAN SYNDROME
CC (SBS), AN AUTOSOMAL, DOMINANT MACROTHROMBOCYTOPENIA CHARACTERIZED
CC BY THROMBOCYTOPENIA, GIANT PLATELETS AND LEUCOCYTE INCLUSIONS.
CC -1- DISEASE: DEFECTS IN MYH9 ARE THE CAUSE OF AN AUTOSOMAL DOMINANT
CC FORM OF NONSYNDROMIC SENSORINEURAL DEAFNESS (DFNA17) WHICH IS
CC CHARACTERIZED BY PROGRESSIVE HEARING IMPAIRMENT AND
CC COCHLEOSACCULAR DEGENERATION.
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: 282215; CAB05105.1; -
DR EMBL: M81105; AAA59888.1; -
DR EMBL: M6180; AAA61765.1; -
DR EMBL: M31013; AAA36349.1; -
DR HSSP: P08799; ILVK.
DR MIM: 160775; -
DR MIM: 153640; -
DR MIM: 155100; -
DR MIM: 603622; -
DR MIM: 605249; -
DR InterPro: IPR000048; IQ.

DR InterPro: IPR004009; Myosin_N.
DR InterPro: IPR002928; Myosin_tail.
DR InterPro: IPR002017; Spectrin.
DR InterPro: IPR001609; myosin_head.
DR Pfam: PF00612; IQ; 1.
DR Pfam: PF00063; myosin_head; 1.
DR Pfam: PF02736; Myosin_N; 1.
DR Pfam: PF01576; Myosin_tail; 1.
DR PRINTS: PR00193; MYOSINHEAVY.
DR ProDom: PD000355; myosin_head; 1.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00242; MISC; 1.
DR PROSITE: PSS0096; IQ; 1.
KW Myosin: ATP-binding; Calmodulin-binding; Actin-binding;
KW Coiled coil; Alkylation; Multigene family; Disease mutation;
KW Deafness.
FT DOMAIN 1 778 MYOSIN HEAD-LIKE.
FT DOMAIN 779 808 IQ.
FT DOMAIN 837 1926 COILED COIL (POTENTIAL).
FT NE_BIND 174 181 ATP (POTENTIAL).
FT DOMAIN 654 676 ACTIN-BINDING.
FT MOD_RES 704 704 ALKYLATION (SH-2) (POTENTIAL).
FT MOD_RES 93 93 ALKYLATION (SH-2) (POTENTIAL).
FT VARIANT /FTID-VAR.010791. N -> K (IN MHA).
FT VARIANT /FTID-VAR.010791. R -> C (IN ETNS).
FT VARIANT /FTID-VAR.010792. R -> C (IN ETNS).
FT VARIANT /FTID-VAR.010792. R -> H (IN DFNA17).
FT VARIANT /FTID-VAR.010793. T -> I (IN MHA).
FT VARIANT /FTID-VAR.010794. R -> C (IN SBS).
FT VARIANT /FTID-VAR.010795. D -> H (IN ETNS).
FT VARIANT /FTID-VAR.010796. E -> K (IN MHA).
FT VARIANT 1841 1841 /FTID-VAR.010797. EAL -> RGH (IN REF. 3).
FT CONFLICT 53 55 T -> S (IN REF. 3).
FT CONFLICT 660 660 C -> M (IN REF. 4).
FT CONFLICT 869 869 C -> Y (IN REF. 4).
FT CONFLICT 931 931 KG -> GR (IN REF. 4).
FT CONFLICT 1240 1241 E -> EE (IN REF. 2).
FT CONFLICT 1350 1350 T -> A (IN REF. 2).
FT CONFLICT 1764 1764 S -> G (IN REF. 2).
FT CONFLICT 1771 1771
SQ SEQUENCE 1960 AA: 226531 MW: 588F84BB8C10656F CRC64:

Query Match 13.9%; Score 149; DB 1; Length 1960;
Best Local Similarity 19.7%; Pred. No. 0.33;
Matches 60; Conservative 68; Mismatches 87; Indels 90; Gaps 10;

QY 4 INKLPDLAEEENV-----LDAEF--LKNELDSYKAQLSQK 38
DB 1071 IAEKQOLAKKEELDAALAVEEAAQKNMALKKIRELSQISELOEDLESBRASRNKA 1130
QY 39 DREKRSQAIITLRPLDEERNATVESIQALNKAEMLCSTLKK-----QMKFLQG 89
DB 1131 EKQRKQGEELAKKLELEPLDSTAAQDLRSKRQDEVIILKKTLEEKATHEAQIQEM 1190
QY 90 RQ-----DETQAREAHRLKCKM---KTMF-----QIELLOS-----ORSEV 125
DB 1191 RQKHSQAVEELALQLEBOTKRVKANLEKAKQOTLENGEGLANEVYLQGGKDESHRKRV 1250
QY 126 EEMIRDMGV-----GQSAVEDLAVYCYSLKREYNLKEARATLADRLKKDLVSSRSKL 181
DB 1251 EAQLQELQYKFNQGERVRELADKVTKLOVELDNTVGLLSQSDSKSKSLTKDFSALESQI 1310
QY 182 K-----TLNTELDQAKLELSAOKDLOSADO-----ETTSIR 213
DB 1311 QDTQELLQDEERNRQKLSLTKLQVDEKNSFRQLEEBEAKHNLEKQIATLHAQVADMK 1370
QY 214 KKSDD 218

```

DB      1371 KKMED 1375

RESULT 35
MYH-CHICK
ID MYH-CHICK STANDARD; PRT; 1978 AA.
AC P10587;
DT 01-JUL-1989 (Rel. 11, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin heavy chain, gizzard smooth muscle.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
NCBI_TaxID=9031;
RX MEDLINE=88118918; PubMed=2892941;
RA Yanagisawa M., Hamada Y., Katsuragawa Y., Imamura M., Mikawa T.,
Masaki T.,
"Complete primary structure of vertebrate smooth muscle myosin heavy
chain deduced from its complementary DNA sequence. Implications on
topography and function of myosin.";
J. Mol. Biol. 198;143-157(1987).
[2]
RP REVISIONS.
RA Masaki T.;
RL Submitted (FEB-1989) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE OF 1-203.
RX MEDLINE=88032919; PubMed=3312184;
RA Maizte T., Onishi H., Yajima E., Matsuda G.;
"Amino acid sequence of the amino-terminal 24 kDa fragment of the
heavy chain of chicken gizzard myosin.";
J. Biochem. 102;133-145(1987).
[4]
RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS) OF 1-818, AND REVISIONS 204-215.
RX MEDLINE=98412652; PubMed=9741621;
RA Dominguez R., Freyzo Y., Trybus K.M., Cohen C.;
"Crystal structure of a vertebrate smooth muscle myosin motor domain
and its complex with the essential light chain: visualization of the
pre-power stroke state.";
Cell 94;559-571(1998).
RL
CC -1- FUNCTION: MUSCLE CONTRACTION.
CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PTM: TWO CYSTEINE RESIDUES IN THE S1 DOMAIN ARE SELECTIVELY
CC ALKYLATED AND ARE REQUIRED FOR MYOSIN ATPASE ACTIVITY.
CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATER BE
CC SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 IQ DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL, X06546; CAA29793.1; -
CC PIR, S03166; S03166.
CC PDB, 1BR1; 09-SEP-98.

```

```

DR PDB; 1BR2; 09-SEP-98.
DR PDB; 1BR4; 09-SEP-98.
DR InterPro; IPR000048; IQ.
DR InterPro; IPR004009; Myosin_N.
DR InterPro; IPR002928; Myosin_tail.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF00063; myosin_head; 1.
DR Pfam; PF02736; Myosin_N; 1.
DR Pfam; PF01576; Myosin_tail; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRODOM; PD000355; myosin_head; 1.
DR SMART; SM00015; IQ; 1.
DR SMART; SM00242; MYSC; 1.
DR PROSITE; PS50096; IQ; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW ATP-binding; Methylation; Alkylation; Calmodulin-binding;
KW Multigene family; 3D-structure.
FT INIT MET 0 0
FT DOMAIN 1 790 MYOSIN HEAD-LIKE.
FT DOMAIN 791 820 IQ.
FT DOMAIN 849 1978 RODLIKE TAIL (S2 AND IMM DOMAINS).
FT DOMAIN 849 1978 COILED COIL (POTENTIAL).
FT NP_BIND 176 183 ATP.
FT DOMAIN 666 688 ACTIN-BINDING.
FT DOMAIN 767 781 ACTIN-BINDING.
FT MOD_RES 1 1 BLOCKED.
FT MOD_RES 127 127 METHYLATION (TRI-) (POTENTIAL).
FT MOD_RES 706 706 ALKYLATION (SH-1).
FT MOD_RES 716 716 ALKYLATION (SH-2).
FT CONFLICT 127 127 MISSING (IN REF. 3).
FT CONFLICT 204 215 KDTSTGGSFS -> RTFASLKVHLP (IN REF. 1).
SQ SEQUENCE 1978 AA; 228663 MW; B7B6C9235E273D93 CRC64;

```

```

Query Match 13.9%; Score 149; DB 1; Length 1978;
Best local similarity 24.3%; Pred. No. 0.33;
Matches 54; Conservative 48; Mismatches 86; Indels 34; Gaps 7;

OY 14 EENNVDAEF-----LKNELDSVKAQLSQDRKRSQAIIIDRLDLEERNATVESLQN 68
DB 892 EEKNLLQEKRLQAEITELTLYAEAEKRVRLAKKQE-----LEELHMEARIEE 941
OY 69 ALKKAEMLCSTLKKQKFFLEQRODETKQAREEHLKC-----KMTQMEQLLELQSO--- 121
DB 942 RSQQLQAEKKRKKQQLMDLEEEAARQKLQLEKVTADGKIKKEDDILIMEDQNNK 1001
OY 122 ---RSEVEEMIRDMGQSAVEQLAVYCVSLKKEYENLKEARKATGELADRLKDLVSS 177
DB 1002 LTKERKLEERVDLTNNLAEEEEKAKNLTIKKNKHSM-----ISELEVRLLKE-EKS 1054
OY 178 RSKIKTLNTELDAQLELRSQKDLQSDQDEI-TSLRKSD 218
DB 1055 RQLEKIKRKLKESSDLHQIAELQAIQIAELAKQAKKEE 1096

RESULT 36
MYSU_RABIT
ID MYSU_RABIT STANDARD; PRT; 501 AA.
AC Q99105;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Myosin heavy chain, embryonic smooth muscle isoform (Fragment).
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RX [1]
RX SEQUENCE FROM N.A.
RX TISSUE=Fetal aorta;
RX MEDLINE=91139672; PubMed=1995631;

```

RA	Kuro-o M,Nagai R., Nakahara K.I., Katoh H., Tsai R.C.,
RA	Tsuchimoto H., Yazaki Y., Ohkubo A., Takaku F.,
RT	"cDNA cloning of a myosin heavy chain isoform in embryonic smooth
RT	muscle and its expression during vascular development and in
RT	arteriosclerosis.";
RL	J. Biol. Chem. 266:3768-3773(1991).
CC	-1- FUNCTION: MUSCLE CONTRACTION.
CC	-1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
CC	HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
CC	AND 2 REGULATORY LIGHT CHAIN SUBUNITS (MLC-2).
CC	-1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC	-1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC	CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC	CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC	-1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC	MEROMYOXIN (LMW) AND 1 HEAVY MEROMYOXIN (HMW). IT CAN LATER BE
CC	SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC	SUBFRAGMENT (S2).
CC	-1- SIMILARITY: THE PERIODICITIES OF HYDROPHOBIC AND CHARGED RESIDUES,
CC	WHICH DICTATE THE ALPHA-HELICAL COILED-COIL STRUCTURE ARE
CC	CONSERVED.
CC	-----
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL: D10280; BAA01124.1; .
DR	PIR: A38650; A38650.
DR	InterPro: IPR002928; Myosin_tail.
DR	Pfam: PF01576; Myosin_tail.1.
KW	Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW	ATP-binding; Multigene family.
FT	NON_TER 1
FT	DOMAIN <1 501 RODLIKE TAIL (S2 AND LMW DOMAINS).
FT	DOMAIN 1 457 COILED COIL (POTENTIAL);
SO	SEQUENCE 501 AA; 58118 MW; 49F793247D00973E CRC64;
OY	Query Match 13.8%; Score 148.5; DB 1; Length 501;
OY	Best Local Similarity 24.5%; Pred. No. 0.088;
OY	Matches 52; Conservative 49; Mismatches 88; Indels 23; Gaps 4;
DB	13 OEENVVLDAEFLNKLNDLVKAQLSQKDREKRQSAIIPDLRTLEERNAYESLGNALNK 72
DB	I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
DB	30 ONKQLRAMEDLMSSKDDVGKNVHELEKSRALEDQVEEMKQTLEDELDGATDALR 89
OY	73 AEMLCTLLKKQMKLFLEGRDE-----TKQAEEAHRLCKMK-----TWEGJEL 116
DB	I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
DB	90 LEVTNQAKAAGEPDLQARDGESEKKRILTFQVRELAELEADERKQALAVASKKKMEI 149
OY	117 LLQSRSVEEKEIRPMGVGSVDEDVAAYCVSLKREYENLKFAKATDELADRLKKDLVS 176
DB	I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
OY	150 DLKDEAOIEAANK---ARRRRVKQLRRLOAQMKRYORELEEARSGSRDEIFAQSKKE---- 202
DB	177 SRSKLTNTLELDAQLELRSAOKLOSADQE 208
DB	I : : : : : I : : : : : I : : : : : I : : : : : I : : : : : I : : : : :
DB	203 SEKKLSLEAFILQLQLEELASSERRARRADE 234
RESULT 37	
MPPL_ARATH	STANDARD: PRT: 727 AA.
ID MPPL_ARATH	
AC Q9LMB5:	
DT 16-OCT-2001 (Rel. 40, Created)	
DT 16-OCT-2001 (Rel. 40, Last sequence update)	
DT 16-OCT-2001 (Rel. 40, Last annotation update)	
DE MAP binding filament-like protein 1.	
GN MPPL OR AT3G16000 OR MS1L.4.	
OS Arabidopsis thaliana (Mouse-ear cress).	

```
CC Eukaryota,Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurostids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI_taxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A..  
RC STRAIN=cv. Columbia:  
RX MEDLINE=20277480; PubMed=10619329;  
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence  
RT features of the regions of 4,504,864 bp covered by sixty pl and TAC  
RT clones.";  
RL DNA Res. 7:131-135(2000).  
CC -I- FUNCTION: BIRDS DNA. INTERACTS WITH CHROMATIN VIA MATRIX  
CC ATTACHMENT REGIONS (MARS). LIKELY TO PARTICIPATE IN NUCLEAR  
CC ARCHITECTURE BY CONNECTING CHROMATIN WITH THE NUCLEAR MATRIX AND  
CC POTENTIALLY WITH THE NUCLEAR ENVELOPE (BY SIMILARITY).  
CC -I- SUBCELLULAR LOCATION: NUCLEAR MATRIX (BY SIMILARITY).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC or send an email to license@sib-sib.ch).  
-----  
DR EMBL; AB012347; BABO2666.1; .  
KW Nuclear protein; DNA-binding; Coiled coil.  
FT DOMAIN 140 692 COILED COIL (POTENTIAL).  
FT DOMAIN 22 35 POLY-SER  
SQ SEQUENCE 727 AA; 82071 MW; 5FAF9A0FF623CAA CRC64;
```

Query Match 13.8%; Score 148.5; DB 1; Length 727;
Best Local Similarity 21.7%; Pred. No. 0.13;
Matches 55; Conservative 53; Mismatches 84; Indels 61; Gaps 8;

Oy	13	QEEENVLDAEFLK-----NELDVSVAKOALSOGRDKRKSQAIIIDLRTDLTEERNAYVES	65
Dd	187	EQEKEKKVEEKAKEBQLSLINOLNSAKDLVETGLRELSSEKKLCBKLRKD-----QIES	239
Oy	66	LONALKNA-----EMLCSTLLRKQMKLEQROD-----ETKRAREBARLKCKM-----	108
Dd	240	LENLSISKAGEDEKALETJKIREKLDIIVEGILODIRNILLSELKOSBEKAAOFNFASLAKKKEAE	299
Oy	109	-----KTMEQJELLLOSOREEVEENIKRMGVGAVSDVEDLAIVCVSIKKEYENTL	156
Dd	300	LKELNSTYTQTSSRDIAEAKELETKQOKKEELIRQSDELDSKNASAIIEELNRIITTLVAEKESY	359
Oy	157	KR-----ARRKATGEADRLRLKKDIVSR-SKLKLTNTLELDQAKLELSAQORDLOSA	205
Dd	360	IQKLDSISKDYSAFKLTSETQAAAADAELISRKEOFIQOLNENLDR-----ALDDVNKS	412
Oy	206	DOETSLLARKKSPD 218	
Dd	413	KDVADVADLTEKYED 425	

RESULT_38
XCP_C_XENILA STANDARD: PRT. 1290 AA.

ID	XCP_C_XENILA	PRT. 1290 AA.
AC	P50532;	
DT	01-OCT-1996 (Rel. 34, Created)	
DN	01-OCT-1996 (Rel. 34, Last sequence update)	
PT	15-JUL-1998 (Rel. 36, Last annotation update)	
DE	Chromosome assembly protein XCAP-C.	
GN	XCAP-C.	
OS	Xenopus laevis (African clawed frog).	
OC	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;	
OC	Xenopodinae; Xenopus.	
OX	NCBI_taxID=8355;	

